Skill and Profit in Active Management

by

Robert F. Stambaugh*

November 14, 2020

Abstract

I analyze skill's role in active management under general equilibrium with many assets and costly trading. More-skilled managers produce larger expected total investment profits, and their portfolio weights correlate more highly with assets' future returns. Becoming more skilled, however, can reduce a manager's expected profit if enough other managers also become more skilled. The greater skill allows those managers to identify profit opportunities more accurately, but active management in aggregate then corrects prices more, shrinking the profits those opportunities offer. The latter effect can dominate in a setting consistent with numerous empirical properties of active management and stock returns.

JEL Classifications: G14, G23

Keywords: active management, investment profit, stock selection

---

*The Wharton School, University of Pennsylvania and NBER. I am grateful for comments from anonymous referees as well as Jonathan Berk, Philip Bond, Vincent Glode, Ron Kaniel, Ľuboš Pástor, Lasse Pedersen, participants at the LSE/FCA/SEBI "Paying for Efficient and Effective Markets" conference, the meeting of the French Finance Association, and seminars at Baruch College, Emory University, Federal Reserve Bank of Atlanta, University of Kansas, University of Pennsylvania, and University of South Carolina.

Skill and Profit in Active Management

Abstract

I analyze skill's role in active management under general equilibrium with many assets and costly trading. More-skilled managers produce larger expected total investment profits, and their portfolio weights correlate more highly with assets' future returns. Becoming more skilled, however, can reduce a manager's expected profit if enough other managers also become more skilled. The greater skill allows those managers to identify profit opportunities more accurately, but active management in aggregate then corrects prices more, shrinking the profits those opportunities offer. The latter effect can dominate in a setting consistent with numerous empirical properties of active management and stock returns.

1. Introduction

Active investment managers seek to outperform a passive benchmark by identifying and trading securities likely to over- or under-perform. The ability to select such securities, i.e., the ability to identify alphas, reflects a manager's skill. The total dollar amount of a manager's outperformance, net of trading costs, is her investment profit. Should more-skilled managers produce greater expected profit than less-skilled managers? If a manager becomes more skilled, does she then produce greater expected profit?

I explore such questions in the context of a general equilibrium model. A large cross-section of managers with different amounts of skill select among many assets and incur costs to trade them. The alphas that managers exploit arise from mispricing induced by "noise" traders. In equilibrium, active management's aggregate asset demands correct the mispricing, but not all of it, and managers produce positive expected investment profits.

Within the cross-section of managers, the model implies that those with more skill produce greater expected profit. Suppose the cross-section shifts, so that some or all managers become more skilled. Should a manager who becomes more skilled then produce greater expected profit? Yes, but not if too many other managers also become more skilled. In that case, becoming more skilled can result in lower expected profit due to the effects on equilibrium asset prices. For example, the managers who become more skilled are more likely to identify correctly a positive-alpha stock and to over-weight it in their portfolios. This collective higher demand for the stock raises the equilibrium price all managers pay for it, reducing the investment profit any manager makes from buying it. This latter effect can dominate and thereby reduce profits even for the managers who become more skilled. They more accurately identify a positive-alpha stock, but the magnitude of the alpha shrinks due to greater price correction. Naturally these same opposing effects of increased skill also apply to negative-alpha stocks, raising the chances such stocks are accurately identified but reducing the investment profits earned by under-weighting or shorting them.

A quantitative specification of the model captures numerous observed characteristics of both the active management industry and asset returns. In that empirically relevant setting, the model implies that if, say, 25% or 50% of managers become more skilled, they earn greater expected profit as a result. Their greater accuracy in identifying alphas outweighs these managers' stronger contribution to price correction. In contrast, if 90% of managers become more skilled, then all managers earn lower expected profit, as the effect of greater price correction then dominates. Of course, in all scenarios, the managers whose skill does not increase earn lower expected profit, because the effect of any increased price correction is not mitigated by an increased ability to identify alphas accurately. What is the effect on the total investment profit of the active management industry when skill increases for some or all managers? Under the same quantitative specification, the model implies that the industry's total expected profit falls even for the above cases in which expected profits rise for the managers who become more skilled.

A manager's investment profits are presumably the principal source of the fund's fee revenue. In the model of Berk and Green (2004), for example, expected investment profit simply equals total fee revenue, as the fund's investors earn zero net alpha in equilibrium. Under the latter condition, therefore, all of the above statements about expected profit apply to fee revenue. I do not impose the zero-alpha condition, but I do assume the manager's objective is to maximize expected profit. No doubt this objective is more easily motivated, the stronger is the manager's perceived association between average profits and fee revenue. Berk and van Binsbergen (2015) document the strong empirical association across funds between total fee revenue and average future investment profits. The authors define investment profit as "value added," and I adopt that term as well later in the exposition. They also propose value added as measuring skill, and the model presented here justifies that interpretation.

The model implies that, within the cross-section of managers, expected profit is determined by skill. A manager takes active (overweight and underweight) positions up to an optimal point at which trading to take larger such positions becomes too costly. A less-skilled manager takes less profitable active positions in reaching that optimum. Each manager reaches such a point, because the cost of trading any asset is convex in the amount of the asset the manager trades. A manager's maximized expected dollar profit does not depend on the dollar amount of the fund's assets under management, AUM. Therefore, funds exhibit decreasing returns to scale, with expected profit per dollar of AUM decreasing in AUM. The assumption that a fund faces decreasing returns to scale with respect to its AUM is central to the model of Berk and Green (2004). The model presented here provides a micro-foundation for that property.

Including trading costs in a model of the active management industry seems desirable, given empirical evidence of such costs' economic importance. For example, Edelen, Evans, and Kadlec (2007) conclude that trading costs present an important source of scale diseconomies for mutual funds. Edelen, Evans, and Kadlec (2013) find that mutual funds' trading costs as a fraction of net asset value are comparable in magnitude to funds' expense ratios. In this respect the model of active management here, with convex trading costs paid to intermediaries, differs from those of Garcia and Vanden (2009) and Gârleanu and Pedersen (2018), for example, in which managers do not face such liquidity costs and the resulting decreasing returns to scale.

In addition to the above insights, the model delivers a number of empirical implications. For example, it predicts a fund's expected profit equals its trading costs. If fee revenue coincides with expected profit, than this prediction is largely consistent with the Edelen, Evans, and Kadlec (2013) evidence that trading costs are similar in magnitude to expense ratios (of which management fees are a substantial part). The model also predicts that, other things equal, a fund charging a higher expense ratio has lower AUM and has portfolio weights that depart more from benchmark weights. Both implications are consistent with evidence reported by Pástor, Stambaugh, and Taylor (2020). An additional implication for future empirical investigation is that a fund with higher average investment profit has a higher correlation across assets between the fund's active weights and individual assets' future returns.

This study's contribution in the context of the previous literature is discussed further in a later section. Included there is a parallel to Grossman and Stiglitz (1980), who identify opposing effects of variance in the signal about the risky asset's payoff. Those opposing effects are similar to the opposing effects of active managers' skill analyzed here.

2. Model assumptions

The model considers a single investment period. Active managers identify and exploit opportunities to outperform the market benchmark by having their funds' weights on individual assets, generally taken to be stocks, deviate from the weights in the market benchmark. The number of active managers, M, is assumed to be large enough for each manager to take stock prices as given in a competitive equilibrium. For active management in aggregate to outperform the market benchmark, before costs and fees, other market participants must underperform that benchmark (e.g., Sharpe (1991)). The latter participants comprise noise traders whose exogenous asset demands are inconsistent with fair values. Black (1986), for example, characterizes noise traders as trading to their own detriment. The noise traders own fraction h of total stock-market wealth and invest in stocks directly. Investors own the remaining fraction $1 - h$ of stock-market wealth. They invest in the stock market through professionally managed funds, either active or passive. Passive managers offer zero-cost market-index funds that invest any money the investors do not allocate to active managers.

2.1. Stocks and alphas

The stock market comprises N stocks, and the supply of each stock equals one share. Stock i has share price p_i at the beginning of the investment period and payoff (including dividends) equal to x_i at the end of the period. It is also useful to define a stock's beginning-of-period "fair value" in order to introduce concepts such as mispricing and price correction. The fair value of stock i is defined as $\bar{p}_i = (1/\mu)\mathrm{E}\{x_i|\Phi\}$, where Φ denotes the union across managers of all signals on which they condition their expectations about x_i, and μ equals one plus the fair-value discount rate, assumed to be exogenous and equal across stocks. There are N shares in the market portfolio, each with price $p_m = (1/N)\sum_{i=1}^{N} p_i$. Any individual stock's p_i can deviate from its $\bar{p}_i$, but the number of stocks, N, is assumed to be large, with deviations between p_i and $\bar{p}_i$ averaging out across stocks. That is, the overall stock market is assumed to be fairly priced, so that p_m is also equal to $(1/N)\sum_{i=1}^{N} \bar{p}_i$. In essence, this assumption means active managers cannot profit from market timing.

Define stock i's market-adjusted return as $r_i = R_i - R_m$, where R_i is the realized rate of return on stock i over the period, and R_m is the return on the market portfolio. Note from the above assumptions that $\mathrm{E}\{1+R_i|\Phi\} = (1/p_i)\mathrm{E}\{x_i|\Phi\} = \mu\,(\bar{p}_i/p_i)$, and $\mathrm{E}\{1+R_m|\Phi\} = \mu$. Define the "alpha" on stock i as $\mathrm{E}\{r_i|\Phi\}$, which is then given by

$$\alpha_i = \mu(\bar{p}_i - p_i)/p_i. \tag{1}$$

That is, α_i is proportional to the fraction of stock i's price represented by mispricing.

2.2. Trading costs

Active managers incur trading costs when taking positions that deviate from market benchmark weights. I treat each fund as already holding the benchmark when entering the model's current period, having previously incurred the sunk cost of initially investing the fund's assets under management (AUM). A manager who simply maintains that benchmark portfolio, as would a passive manager, incurs no trading costs in the current period, whereas taking active positions that depart from benchmark weights incurs trading costs. Define manager j's active weight in stock i as

$$\phi_i^{(j)} = \phi_{A,i}^{(j)} - \phi_{m,i}, \tag{2}$$

where $\phi_{A,i}^{(j)}$ is stock i's weight in manager j's portfolio, and $\phi_{m,i}$ is stock i's weight in the market portfolio. Note that $\sum_{i=1}^{N} \phi_i^{(j)} = 0$. The dollar amount of stock i traded by manager

j is assumed to be

$$D_i^{(j)} = \phi_i^{(j)} W^{(j)}, \tag{3}$$

with $W^{(j)}$ denoting fund j's dollar AUM.

The $\phi_i^{(j)}$'s reflect the discretionary active trades that managers make with their existing AUM. Excluded are costs of investing or disinvesting any flows into or out of the fund. In this respect, the single period is envisioned as one in the midst of many for an ongoing fund with no significant flows in or out. In the model of Berk and Green (2004), for example, flows to and from funds occur because investors update assessments of fund managers' skills. One might thus view the current model as abstracting from investors' learning about skill, essentially treating such learning as having already occurred. Any active manager's investment activity (excluding inflows and outflows) can be represented as placing the fund's total AUM in a buy-and-hold long position in the market portfolio and then actively trading a zero-investment portfolio of long and short positions in individual stocks.[1] The trading costs modeled here pertain to the latter portfolio and are incurred in any period that is viewed as the model's single period. The one-time cost of establishing the buy-and-hold market position is excluded, meaning that if a fund were to behave as an index fund, the model would assign it zero trading costs.

In the quantitative analysis presented later, a single period is taken as a year. In reality, which is of course multiperiod, active positions can span periods. If, say, manager j gives stock i positive $\phi_i^{(j)}$'s across two periods, then the fund's trade in the second period is less than $D_i^{(j)}$ in equation (3). Purely for tractability, I abstract from that multiperiod reality when casting the model within a single period. The basic concept underlying equation (3) is simply that active managers make bets each period, underweighting and overweighting stocks relative to the market, and generally the managers trade to place those bets.

I assume the trading costs faced by active managers are convex in the amount traded. Specifically, for any manager j, the round-trip dollar cost $C_i^{(j)}$ of trading dollar amount $D_i^{(j)}$ of stock i is

$$C_i^{(j)} = c d_i^{(j)} D_i^{(j)}, \tag{4}$$

where c is a positive constant and $d_i^{(j)}$ is the fraction of stock i's total market capitalization represented by $D_i^{(j)}$. That proportional costs increase with trade size is a familiar concept,

[1]Note that this representation need not conflict with, say, being a long-only fund. A short position for stock i in the zero-investment long-short portfolio, i.e., $\phi_i^{(j)} < 0$, does not imply the fund has an actual short position in the stock, which occurs when $\phi_i^{(j)} < -\phi_{m,i}$, where $\phi_{m,i}$ is stock i's market-portfolio weight. The latter case, if it occurs, is not treated differently here, again for tractability. In reality, taking actual short positions can be costlier than taking long positions.

well supported by evidence, e.g., Keim and Madhavan (1997). The specification in equation (4), which is also assumed by Stambaugh (2014) and Pástor, Stambaugh, and Taylor (2020), has the proportional trading cost increase linearly with trade size. A linear function for a stock's proportional trading cost is entertained, for example, by Kyle and Obizhaeva (2016). That study examines portfolio transition trades and concludes that a linear function fits the data only slightly less well than a nonlinear square-root specification. A linear specification is imposed here, again for tractability. With the proportional trading cost being linear in the amount traded, $D_i^{(j)}$, the total trading cost, $C_i^{(j)}$, is quadratic in $D_i^{(j)}$. (Note that $d_i^{(j)} = D_i^{(j)}/p_i$ in equation (4).) Quadratic trading costs are also assumed, for example, in the models of Heaton and Lucas (1996), Grinold (2006), and Gârleanu and Pedersen (2013).

The trading cost $C_i^{(j)}$ in equation (4) does not include the impact of the manager's demand for the stock on the equilibrium price, p_i. The aggregate demand for the stock across the model's many managers impacts p_i, but I assume a competitive equilibrium in which each manager views p_i as given, unaffected by her own relatively infinitesimal demand. The $C_i^{(j)}$'s instead represent compensation to liquidity-providing intermediaries (broker-dealers) who take short-lived positions that simply facilitate ultimate market clearing among funds and noise traders. One can envision many intermediaries accessing different sources of liquidity or acting at slightly different times, such that the above trading costs represent very transitory price impacts, essentially brief excursions to effective bid or ask prices that accompany the price-formation process at the start of the model's single period. These trading costs, which constitute revenue to the intermediaries, might best be viewed as the traditional inventory-related costs of market-making reflected in effective bid-ask spreads (e.g., Stoll (1979) and Amihud and Mendelson (1980)). Intermediaries play no role in the model other than to receive the $C_i^{(j)}$'s from managers in exchange for providing liquidity. A bid-ask spread's adverse-selection component that provides zero expected profit to intermediaries (e.g., Glosten and Milgrom (1985)), rather than being part of $C_i^{(j)}$, is better viewed as another unmodeled element of market microstructure underlying the formation of the equilibrium price, p_i.

2.3. Noise traders

Let $\phi_{H,i}$ denote the weight on stock i in the aggregate portfolio of the noise traders, with $\phi_{H,i}$ assumed to be exogenous and non-negative (no short selling by the noise traders). The price of stock i that would arise in an equilibrium with no active management, with prices thus determined solely by the asset demands of noise traders, is equal to[2]

$$\hat{p}_i = Np_m\phi_{H,i}. \tag{5}$$

I assume

$$\hat{p}_i = v_i\bar{p}_i, \tag{6}$$

with v_i (≥ 0) independently and identically distributed (i.i.d.) across stocks (i), independent of $\bar{p}_i$, and with

$$\mathrm{E}\{v_i\} = 1. \tag{7}$$

In essence, the "noise" in noise-trader demands would distort prices away from fair values, and those relative distortions are i.i.d. across stocks and zero on average.

2.4. Skill and forecasts

The manager of fund j receives a set of N signals, $s_1^{(j)}, \ldots, s_N^{(j)}$, with

$$s_i^{(j)} = \alpha_i + \eta_i^{(j)}, \tag{8}$$

where $\eta_i^{(j)}$ has zero mean, is independent of α_i, and has variance equal to $\sigma_\eta^{2(j)}$. The $\eta_i^{(j)}$'s are also independent across both stocks and funds.[3] One interpretation of these signals is that manager j has a team of analysts who supply her with payoff-relevant information about a stock summarized by $s_i^{(j)}$. The analysts observe equilibrium prices, the p_i's, but nevertheless assess the x_i's (and thus the $\bar{p}_i$'s and α_i's) differently from other analysts. I do not specify the mechanism underlying these differences. Banerjee (2011), for example, summarizes various explanations for disagreement under both rational-expectations and difference-in-opinion approaches. One possibility is simply that the analysts lack sufficient skill to extract all of the information embedded in equilibrium prices, so that the $\eta_i^{(j)}$'s reflect these limitations. In this narrative, the fund manager relies on her analysts in the sense that she does not extract payoff-relevant information from prices beyond what is reflected in the $s_i^{(j)}$'s she receives.

The fund manager does, however, have rational expectations about the usefulness of her team's $s_i^{(j)}$'s for predicting returns, presumably drawing on her experience. Specifically, she

[2]Without active management, investors place all their stock-market wealth in an index fund, so their portfolio's weight on stock i is $\phi_{S,i} = \phi_{m,i} = p_i/(Np_m)$. Combining the latter with the market-clearing condition $(1-h)\phi_{S,i} + h\phi_{H,i} = \phi_{m,i}$ gives $p_i = Np_m\phi_{H,i}$, which is the right-hand side of equation (5).

[3]The $\eta_i^{(j)}$'s cannot literally be independent across j and independent of α_i if the latter, as defined earlier, is the conditional expectation of r_i given $\Phi = \{s_i^{(1)}, \ldots, s_i^{(M)}\}$, but this technicality seems reasonably ignored when M is large, as in the model's quantitative applications later.

computes $\tilde{\alpha}_i^{(j)} = \mathrm{E}\{r_i | s_1^{(j)}, \ldots, s_N^{(j)}\}$, $i = 1, \ldots, N$. Given the independence assumptions about the $\eta_i^{(j)}$'s, the signal relevant to stock i is simply $s_i^{(j)}$. In addition, I assume the manager conditions linearly on that signal, thereby computing the forecasts,

$$\tilde{\alpha}_i^{(j)} = \lambda_i^{(j)} s_i^{(j)}, \quad i = 1, \ldots, N, \tag{9}$$

with

$$\lambda_i^{(j)} = \frac{\mathrm{Cov}\{r_i, s_i^{(j)}\}}{\mathrm{Var}\{s_i^{(j)}\}}. \tag{10}$$

Note that because $s_i^{(j)}$ is uncorrelated with $r_i - \alpha_i$ and $\eta_i^{(j)}$ is uncorrelated with α_i,

$$\lambda_i^{(j)} = \frac{\sigma_{\alpha,i}^2}{\sigma_{\alpha,i}^2 + \sigma_\eta^{2(j)}}, \tag{11}$$

with $\sigma_{\alpha,i}^2$ denoting the unconditional variance of α_i. The smaller is a manager's $\sigma_\eta^{(j)}$, the more she weights the signals she receives, and the greater is her resulting skill to identify the α_i's accurately. Throughout I equate lower $\sigma_\eta^{(j)}$ to greater skill.

The simple specification in equation (9), assumed for tractability, imposes linear conditioning and a zero intercept. In general, neither property obtains formally here from the joint distribution of $s_i^{(j)}$ and the equilibrium α_i. As demonstrated later, however, both properties can be close approximations in the context of the calibrated model. The number of assets, N, is taken as large, such that the distributions of the v_i's and $\eta_i^{(j)}$'s across i correspond to the continuous distributions characterizing the randomness in those quantities for a given i. Distributions of the v_i's and $\eta_i^{(j)}$'s can be specified, for example, to give $\mathrm{E}\{\alpha_i | s_i^{(j)}\}$ a Pearson correlation of 0.998 with $s_i^{(j)}$ across the individual percentiles of $s_i^{(j)}$ and to give α_i a nearly zero mean (2 basis points per annum).

Given many (thousands) of managers, I assume that α_i in equation (8) is well approximated by the weighted average of the $s_i^{(j)}$'s across managers, with the $\lambda^{(j)}$'s as weights, so that,

$$(1/M) \sum_{j=1}^{M} \left(\frac{\lambda_i^{(j)}}{\lambda_i} \right) \eta_i^{(j)} = 0, \tag{12}$$

with $\lambda_i = (1/M) \sum_{j=1}^{M} \lambda_i^{(j)}$. Given the independence assumptions for the $\eta_i^{(j)}$'s, the earlier definition of α_i as $\mathrm{E}\{r_i | \Phi\}$ requires $\alpha_i = \mathrm{E}\{r_i | s_i^{(1)}, \ldots, s_i^{(M)}\}$. The same independence assumptions imply that, as M grows large, this conditional expectation of r_i approaches a weighted average of the $s_i^{(j)}$'s across managers, $(1/M) \sum_{j=1}^{M} (\overline{\sigma_\eta^2} / \sigma_\eta^{2(j)}) s_i^{(j)}$, with $\overline{\sigma_\eta^2}$ denoting the harmonic mean of $\sigma_\eta^{2(j)}$, equal to $\left[(1/M) \sum_{j=1}^{M} 1/\sigma_\eta^{2(j)}\right]^{-1}$. The equivalent relation in terms of the $\eta_i^{(j)}$'s is that their same weighted average is zero,

$$(1/M)\sum_{j=1}^{M}\left(\frac{\overline{\sigma_\eta^2}}{\sigma_\eta^{2(j)}}\right)\eta_i^{(j)} = 0. \tag{13}$$

From equations (8) through (11), equation (12) is equivalent to the condition

$$(1/M)\sum_{j=1}^{M}\left(\frac{\overline{\sigma_{\alpha,i}^2+\sigma_\eta^2}}{\sigma_{\alpha,i}^2+\sigma_\eta^{2(j)}}\right)\eta_i^{(j)} = 0, \tag{14}$$

with $\overline{\sigma_{\alpha,i}^2+\sigma_\eta^2}$ denoting the harmonic mean of $\sigma_{\alpha,i}^2+\sigma_\eta^{2(j)}$. If $\sigma_\eta^{2(j)}$ is identical across managers (as in an some of the later calibrations), then of course equations (13) and (14) are equivalent. When $\sigma_\eta^{2(j)}$ differs across managers, the difference in weightings depends on the size of $\sigma_{\alpha,i}^2$ in equilibrium relative to the typical $\sigma_\eta^{2(j)}$. In the later calibrations, $\sigma_\eta^{2(j)}$ is typically at least an order of magnitude larger than $\sigma_{\alpha,i}^2$. I impose equation (12) for tractability in deriving equilibrium prices, but it seems reasonable to assume that independent signal noise generally averages out across the large number of managers, especially when the noisier signals receive less weight, as they do in both equations (13) and (14).

2.5. Managers' objective

Berk and van Binsbergen (2015) define the fund's value added as its AUM times its benchmark-adjusted rate of return net of trading costs but before fees charged to investors. Denote that return for manager j as $r_A^{(j)}$, so that the manager's value added, her investment profit, is therefore

$$V_A^{(j)} = r_A^{(j)}W^{(j)}. \tag{15}$$

Let $F^{(j)}$ denote fund j's total fee revenue, paid by investors as a deduction from the investment profit, and define $\psi^{(j)} = F^{(j)}/\mathrm{E}\{V_A^{(j)}\}$. In the model of Berk and Green (2004), for example, a fund's expected value added is equal to its fee revenue, because in equilibrium the fund's expected benchmark-adjusted return net of both trading costs and fees is equal to zero. This zero-net-alpha condition is equivalent to $\psi^{(j)} = 1$. A somewhat less restrictive view is that managers see producing higher average value added as the path to higher fee revenue, essentially viewing $\psi^{(j)}$ as positive and relatively fixed but not necessarily equal to 1. Accordingly, I simply assume that managers seek the highest expected value added, with the implicit motivation being maximizing fee revenue.

Specifically, manager j's objective is to choose the fund's active weights to maximize the fund's expected value added, conditional on her signals and chosen weights. In performing this maximization, the manager takes the fund's current AUM, $W^{(j)}$, as given, so the equivalent objective is

$$\max_{\phi_1^{(j)},\ldots,\phi_N^{(j)}} \mathrm{E}\left\{r_A^{(j)}|s_1^{(j)},\ldots,s_N^{(j)},\phi_1^{(j)},\ldots,\phi_N^{(j)}\right\}, \quad \text{subject to} \sum_{i=1}^N \phi_i^{(j)} = 0. \tag{16}$$

The fund's AUM likely depends on its fee structure, with, say, a higher fixed fee rate simply implying a lower AUM, other things equal. As will be shown, however, the fund's maximized expected value added does not depend on $W^{(j)}$ and thus not on the fee structure, so the latter is not relevant to the general model. Later, in calibrating the model, I assume a fixed proportional fee rate but do not impose a zero-net-alpha restriction that equates fee revenue to expected value added.

With the manager assumed to maximize expected $r_A^{(j)}$, the fund is essentially risk neutral with respect to its benchmark-adjusted return. This property coincides, for example, with that of the fund's investors in the Berk and Green (2004) model. Those investors do not require expected compensation (net alpha) for bearing risk in the fund's benchmark-adjusted return.

2.6. Equilibrium conditions

An equilibrium occurs when each fund's active weights $(\phi_1^{(j)},\ldots,\phi_N^{(j)})$ satisfy the objective in equation (16) and stock prices $(p_1,\ldots,p_N)$ satisfy the market-clearing conditions

$$h\phi_{H,i} + (1-h)\phi_{S,i} = \phi_{m,i}, \quad i = 1,\ldots,N, \tag{17}$$

where $\phi_{m,i} = p_i / \sum_{j=1}^N p_j$ is stock i's market weight, and $\phi_{S,i}$ is the stock's weight in investors' aggregate stock portfolio. The latter portfolio combines the aggregate active portfolio with the index fund.

3. Model implications

3.1. Active weights

Proposition 1. *The active weight of fund j in stock i is given by*

$$\phi_i^{(j)} = \frac{p_i \tilde{\alpha}_i^{(j)}}{2cW^{(j)}}. \tag{18}$$

Note that active weights are inversely proportional to AUM, as $W^{(j)}$ appears in the denominator in equation (18). That is, other things equal, a smaller AUM simply implies proportionately larger magnitudes for active weights, such that the dollar active positions in each stock are unaffected. In the fixed-fee-rate, zero-net-alpha scenario discussed earlier, recall that, holding skill constant, a higher fee rate implies a lower AUM. The lower AUM then implies proportionately larger deviations from benchmark weights, by equation (18). This implication is consistent with the empirical evidence of Pástor, Stambaugh, and Taylor (2020), who find that among active equity mutual funds, those with a higher fee rate, proxied by the expense ratio, essentially tend to have portfolio weights that deviate more from benchmark weights.

3.2. Stock prices

Proposition 2. *The equilibrium price of stock i is given by*

$$p_i = \bar{p}_i + \theta(\hat{p}_i - \bar{p}_i), \tag{19}$$

where the positive scalar $\theta \in (0, 1)$ is the solution to

$$\theta = \left(1 + \frac{\lambda \mu M}{2hc}\right)^{-1}, \tag{20}$$

with

$$\lambda = (1/M) \sum_{j=1}^{M} \lambda^{(j)}, \tag{21}$$

in which

$$\lambda^{(j)} = \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_\eta^{2(j)}}, \tag{22}$$

and $\sigma_\alpha^2 = \mathrm{Var}\{\alpha_i\}$, equal across stocks, is given by

$$\sigma_\alpha^2 = \mathrm{E}\{\alpha_i^2\} - (\mathrm{E}\{\alpha_i\})^2, \tag{23}$$

with

$$\alpha_i = -\mu \frac{\theta(v_i - 1)}{1 + \theta(v_i - 1)}. \tag{24}$$

The model inputs consist of the cross-section of skill ($\sigma_\eta^{(j)}$'s), the distribution of noise-trader demands (v_i's), and the scalar parameters M, h, c, and μ. Given θ, μ, and the probability distribution of v_i, the expectations in equation (23) can be computed numerically. The resulting σ_α, along with the cross-section of $\sigma_\eta^{(j)}$'s, gives the value of λ via equations

(21) and (22). That value of λ, along with M, h, and c then implies θ from equation (20). These equations are solved numerically for θ, and in all later calibrations a unique positive solution for θ obtains over a wide range of starting values.

3.3. Aggregate active weights and active share

The aggregate active weight in stock i is its AUM-weighted active weight across funds: $\phi_i = (1/M)\sum_{j=1}^{M}(W^{(j)}/W)\phi_i^{(j)}$, with $W = (1/M)\sum_{j=1}^{M} W^{(j)}$.

Proposition 3. *The aggregate active weight in stock i is given by*

$$\phi_i = \frac{\mu\lambda(\bar{p}_i - p_i)}{2cW}. \tag{25}$$

Aggregate active weights are inversely proportional to average fund AUM, with W appearing in the denominator in equation (25). This relation mirrors the corresponding inverse relation between AUM and active weights for an individual fund. Observe from equation (25) that a stock's aggregate active weight also is proportional to the stock's mispricing, $\bar{p}_i - p_i$, with an overpriced (underpriced) stock receiving a negative (positive) active weight. In other words, active management in aggregate reduces mispricing.

The cross-sectional dispersion in a portfolio's active weights can be summarized by active share, a measure proposed by Cremers and Petajisto (2009). Using their definition, the active share of the aggregate active portfolio equals $AS = (1/2)\sum_{i=1}^{N}|\phi_i|$. The following proposition gives this measure when the number of stocks, N, is large, such that $(1/N)\sum_{i=1}^{N}|\phi_i| = \mathrm{E}\{|\phi_i|\}$.

Proposition 4. *For N large, the aggregate active portfolio's active share is*

$$AS = \frac{M\mu\lambda\theta}{4c\delta(1-h)}\mathrm{E}\{|v_i - 1|\}, \tag{26}$$

where δ, active management's market share, i.e., the aggregate fraction of investors' stock-market wealth that is actively managed, is given by

$$\delta = \frac{\sum_{j=1}^{M} W^{(j)}}{(1-h)Np_m}. \tag{27}$$

The inverse relation between AUM and the magnitudes of active weights noted above also arises for aggregate active share in equation (26). That is, the larger is active management's market share, δ, the lower is its active share, AS. For example, active management could increase δ simply by decreasing its fee rates. With no changes in managers' skill, noise-trader demands, and the other basic parameters in Proposition 2, nothing other than δ changes in equation (26), and thus AS decreases.

3.4. Expected value added

The unconditional expectation of a manager's dollar value added, $\mathrm{E}\{V_A^{(j)}\}$, is a central implication of the model. The following proposition characterizes this quantity, scaled by the total dollar value of the stock market, Np_m, simply to produce a unitless measure.

Proposition 5. *Define* $\Pi^{(j)} = \mathrm{E}\{V_A^{(j)}\}/(Np_m)$. *Then*

$$\Pi^{(j)} = \frac{\tilde{\sigma}_\alpha^2 + \sigma_\eta^{2(j)}}{4c\left(1 + \sigma_\eta^{2(j)}/\sigma_\alpha^2\right)^2}, \tag{28}$$

with

$$\tilde{\sigma}_\alpha^2 = \mathrm{E}\left\{\frac{p_i}{p_m}\alpha_i^2\right\} = (\mu\theta)^2 \mathrm{E}\left\{\frac{(v_i - 1)^2}{1 + \theta(v_i - 1)}\right\}. \tag{29}$$

The expectations in equation (29) are identical across i (assets).

The quantity $\tilde{\sigma}_\alpha^2$ defined in equation (29) can be interpreted as the variance of the α_i's with respect to the value-weighted probability measure, p_i/p_m. Under this measure, the variance is simply the expected α_i^2 because the expected α_i is zero (the value-weighted average alpha is zero by construction). Both this variance, $\tilde{\sigma}_\alpha^2$, and the regular variance, σ_α^2, reflect expected magnitudes in equilibrium α_i's. These related measures thus both summarize equilibrium mispricing in the cross-section of stocks, with larger values interpreted as greater overall mispricing. Observe that $\Pi^{(j)}$ is increasing in both σ_α^2 and $\tilde{\sigma}_\alpha^2$, consistent with greater mispricing allowing any manager to produce greater value added, whatever her skill.

3.5. Skill and value added

Berk and van Binsbergen (2015) argue that managers with greater skill produce greater value added, and they propose $\mathrm{E}\{V_A^{(j)}\}$ as capturing skill differences across managers. The following proposition justifies doing so.

Proposition 6. *Within a given equilibrium,* $\mathrm{E}\{V_A^{(j)}\}$ *is a decreasing function of* $\sigma_\eta^{(j)}$ *for* $\sigma_\eta^{2(j)} > \sigma_\alpha^2 - 2\tilde{\sigma}_\alpha^2$. *The latter inequality holds for all* $\sigma_\eta^{(j)}$ *if* $\theta > 0.5$.

The above condition that $\theta > 0.5$ appears to be very weak. In all of the quantitative analyses presented later, including those departing substantially from a setting capturing numerous empirical features of active management and asset returns, θ is much closer to 0 than to 1. Thus, within settings that seem reasonable to entertain, expected value added is greater for the managers who more accurately identify α_i's, and this statement applies for all levels of such skill. Also important to note, however, is that this statement about skill's relation to value added applies within a given equilibrium, i.e., for a given distribution of $\sigma_\eta^{(j)}$'s across managers. If the $\sigma_\eta^{(j)}$'s of some or all managers change, then so does the amount of equilibrium price correction reflected by σ_α^2 and $\tilde{\sigma}_\alpha^2$ in equation (28).

3.6. Correlations between weights and returns

The following proposition links the cross-section of managers' skills to the correlation between each manager's active weights and future benchmark-adjusted returns.

Proposition 7. *The correlation across stocks between r_i and $\phi_i^{(j)}$ is a decreasing function of $\sigma_\eta^{(j)}$ within the cross-section of managers.*

When combined with Proposition 6, the above proposition provides an implication that can be investigated empirically. That is, funds whose active weights better predict benchmark-adjusted returns should also exhibit greater average value added. This implication can be investigated using holdings of active mutual funds, while recognizing some inherent limitations. A reasonable concession to the degree of abstraction in the simple model would be to test this implication within style sectors (e.g., separately for large-cap funds and small-cap funds). Also, funds' holdings are reported quarterly, obscuring effects of intra-quarter portfolio choices that could be nontrivial, given the evidence of Kazperczyk, Sialm, and Zheng (2008).

A number of previous studies, including Grinblatt and Titman (1993), Daniel, Grinblatt, Titman, and Wermers (1997), and Kacperczyk, van Nieuwerburgh, and Veldkamp (2014), estimate a fund's stock-selection skill as essentially a covariance between weights and future returns. For example, Kacperczyk, van Nieuwerburgh, and Veldkamp (2014) measure stock-picking skill as $\sum_{j=1}^N \phi_i^{(j)} r_i$, which is proportional to the covariance across stocks between r_i and $\phi_i^{(j)}$ (recall $\sum_{j=1}^N \phi_i^{(j)} = 0$).

An important distinction to be drawn, however, is that the covariance between r_i and $\phi_i^{(j)}$ must be divided by the cross-sectional standard deviation of $\phi_i^{(j)}$ to compute the correlation in

Proposition 7. Controlling for this cross-sectional dispersion in a fund's active stock weights is necessary to capture differences in skill across funds. Otherwise, what one identifies as skill differences using just the above covariance is potentially contaminated by differences across funds in fee rates and their accompanying AUMs that are unrelated to skill. Recall from the discussion of Proposition 1 that, holding constant the manager's return predictions, which embody any skill, a higher fee rate and thus lower AUM implies greater magnitudes (and thus a higher standard deviation) of the $\phi_i^{(j)}$'s.

3.7. Trading costs and net alpha

With $C_i^{(j)}$ given in equation (4), fund j's total dollar trading cost is $C^{(j)} = \sum_{i=1}^{N} C_i^{(j)}$. This cost has a simple relation to the fund's value added:

Proposition 8.

$$\mathrm{E}\left\{C^{(j)}\right\} = \mathrm{E}\left\{V_A^{(j)}\right\} \tag{30}$$

The above result can reinterpreted in terms of the fund's net return to investors. Recall that $F^{(j)}$ denotes the fund's total fee revenue, and define the corresponding proportional fee rate $f^{(j)} = F^{(j)}/W^{(j)}$. Also recall that the fund's investment return before fees is $r_A^{(j)} = V_A^{(j)}/W^{(j)}$, so the expected net benchmark-adjusted return for an investor in the fund is $\alpha_{net}^{(j)} = \mathrm{E}\left\{r_A^{(j)}\right\} - f^{(j)}$. Combining these relations with equation (30) gives the following implication about the fund's net alpha.

Corollary 1.

$$\alpha_{net}^{(j)} = \frac{\mathrm{E}\left\{C^{(j)}\right\}}{W^{(j)}} - f^{(j)} \tag{31}$$

Empirical estimates of equity mutual funds' average net alphas are generally small in magnitude (as discussed later), and the model of Berk and Green (2004) predicts zero net alpha in equilibrium. If $\alpha_{net}^{(j)} \approx 0$, then equation (31) implies that $f^{(j)}$ should be close to the fund's trading costs divided by AUM. This implication seems consistent with the evidence of Edelen, Evans, and Kadlec (2013), for example, who observe that estimated trading costs of equity mutual funds as a fraction of AUM are similar in magnitude to funds' expense ratios.

Suppose that a fund charges a fixed proportional fee rate, $f^{(j)}$, and that its expected value added equals fee revenue, as in the zero-net-alpha setting of Berk and Green (2004), so $f^{(j)}W^{(j)} = \mathrm{E}\{V_A^{(j)}\}$. Then charging a higher fee rate simply implies a proportionately lower

AUM, because expected value added, which is determined by skill, does not change. In other words, holding skill constant, higher-fee funds should be smaller. Pástor, Stambaugh, and Taylor (2020) observe empirically a strong negative cross-sectional correlation between fund size and expense ratio, and they predict such a result as long as differences in skill are not large.

Recall that W is the average fund AUM, $\sum_{j=1}^{M} W^{(j)}/M$, and let C denote the average expected trading cost, $\sum_{j=1}^{M} \mathrm{E}\{C^{(j)}\}/M$. The AUM-weighted average across funds of $\mathrm{E}\{C^{(j)}/W^{(j)}\}$ then equals C/W. Also, from equation (31), $\alpha_{net} = C/W - f$, where α_{net} and f denote the AUM-weighted averages of $\alpha_{net}^{(j)}$ and $f^{(j)}$. Recall as well that $MW = \delta(1-h)Np_m$, and that $\Pi^{(j)} = \mathrm{E}\{V_A^{(j)}\}/(Np_m)$. Combining these relations with equation (30) gives the following corollary.

Corollary 2.

$$\alpha_{net} + f \;=\; \frac{C}{W} \;=\; \frac{\sum_{j=1}^{M} \Pi^{(j)}}{\delta(1-h)}. \tag{32}$$

Equation (32) proves useful in the subsequent quantitative analysis. In particular, recall from Proposition 5 that if the $\sigma_\eta^{(j)}$'s are equal across managers, then so are the $\Pi^{(j)}$'s in the numerator of the right-most expression above.

3.8. Decreasing returns to scale

Note that $\mathrm{E}\{r_A^{(j)}\} = \mathrm{E}\{V_A^{(j)}\}/W^{(j)}$. Also, $\mathrm{E}\{V_A^{(j)}\}$ in equilibrium does not depend on the fund's AUM. That is, $W^{(j)}$ does not enter $\Pi^{(j)}$ in equation (28). The following corollary then follows.

Corollary 3. *For a given fund j, $\mathrm{E}\{r_A^{(j)}\}$ is decreasing in $W^{(j)}$, as the product of those quantities is a constant (expected value added).*

Each individual fund therefore faces decreasing returns to scale (DRS): the fund's expected (benchmark-adjusted) investment profit per dollar of AUM is decreasing in total AUM. Recent studies such as Zhu (2018), using fund returns, and Pástor, Stambaugh, and Taylor (2020), using fund characteristics, find evidence of fund-level DRS. The Berk and Green (2004) model assumes each fund faces DRS. The model here has DRS as an implication. If the fund's fee revenue equals its expected value added, as in Berk and Green (2004), then if the fund raises its AUM by lowering its fee rate, the above corollary implies its $\mathrm{E}\{r_A^{(j)}\}$ falls accordingly.

Pástor and Stambaugh (2012) assume a fund has decreasing returns with respect to the scale of its competition, essentially the entire active management industry. The rationale is that more money being deployed actively in aggregate has greater equilibrium price impacts, thereby shrinking alphas and lowering any fund's investment profit. In the current model, such equilibrium pricing effects on each fund's profits occur endogenously.

4. Quantitative analysis

This section explores the model's implications about the effects of skill on expected value added. I specify the model quantitatively in a setting that captures numerous observed properties of both the active management industry and asset returns. Against this backdrop, I analyze the effects on managers' value added when the skill of some managers shifts relative to a benchmark case in which all have equal skill.

4.1. Model inputs and the benchmark setting

Some of the model's inputs have direct empirical analogs. For example, I set μ equal to 1.065, corresponding to one plus the average U.S. stock-market return over recent decades (though results are not sensitive to this parameter). Some other quantities equal approximately their mid-sample values from Stambaugh (2014). Specifically, I set the fraction of stock market wealth owned by noise traders, h, to 0.30; I set active management's market share, δ, equal to 0.8; and I set the number of active managers, M, to 2200. The average fee rate, f, is set to 0.012, equal to the average active equity fund's expense ratio computed by Pástor, Stambaugh, and Taylor (2020) but also close to the quantity's mid-sample trend point of 0.011 from Stambaugh (2014). The remaining model inputs are c, σ_η, and the distribution of v_i. Empirical analogs for these three inputs are more elusive, and I do not attempt a correspondence for each one individually. Instead, as explained below, I jointly specify them to deliver a seemingly reasonable set of additional empirical implications.

I first assume v_i (> 0) obeys a Weibull distribution, which is characterized by two parameters that determine its scale and shape.[4] Because $\mathrm{E}\{v_i\} = 1$ by assumption, there is one free parameter, k, that determines the distribution's shape. Figure 1 displays densities under alternative values of k. As k increases, the density concentrates around $v_i = 1$, corresponding to relatively modest deviations of $\hat{p}_i$ from $\bar{p}_i$ in equation (6). As k decreases, the density's

[4]For a discussion of the Weibull distribution, see Johnson and Kotz (1970, chapter 20).

mass moves toward zero and skewness increases. One can view the latter case as noise traders collectively putting low weights on many stocks and large weights on a relative few. As Stambaugh (2014) discusses, numerous empirical studies indicate that direct holdings of individuals are quite undiversified and exhibit significant commonality across individuals.[5] Commonality in holdings limits the extent to which the low diversification by individuals washes out when their holdings aggregate. To the extent that noise trading reflects direct holdings by individuals, the evidence of such commonality is consistent with low values of k.

I then specify k, c, and σ_η to have the model deliver seemingly reasonable values for a rather diverse set of quantities:

1. active management's active share

2. active management's trading cost as a fraction of AUM

3. active management's net alpha

4. the dispersion in individual stocks' (gross) alphas

The first three are obtained using equations (26) and (32), while the fourth uses the distribution of v_i combined with α_i in equation (24). For ease of discussion and interpretation, k, c, and σ_η are constrained to round values having just one significant digit.

Table 1 summarizes the parameter values and implied quantities. Panel A lists the values given earlier and those specified for c, k, and σ_η. The choice of $c = 2$ means, for example, that trading 10 basis points of a stock's outstanding shares incurs a round-trip cost of 20 basis points, which seems at least plausible given the empirical analysis of institutional trading costs by Keim and Madhavan (1997). For example, that study's regression estimates imply the same round-trip cost for that trade size when a value-oriented manager trades the stock of an NYSE-listed company whose stock has a \$20 share price and a \$1.5 billion market capitalization.[6] The input of $k = 0.6$ is consistent with the previously discussed scenario of relatively low aggregate diversification of noise traders, as illustrated in Figure 1. The input of $\sigma_\eta = 0.06$ means each active manager receives rather imprecise signals about stocks' annual alphas.

[5]Studies presenting evidence of poor diversification include Blume, Crockett, and Friend (1974), Lease, Lewellen, and Schlarbaum (1974), Blume and Friend (1975), Kelly (1995), Polkovnichenko (2005), and Goetzmann and Kumar (2008). Evidence of significant commonality in individuals' stock holdings is reported by Feng and Seasholes (2004), Dorn, Huberman, and Sengmueller (2008), Barber and Odean (2008), and Barber, Odean, and Zhu (2009).

[6]Their regression, whose coefficient estimates are reported in their Table 5, gives the expected one-way trading cost as 10 basis points. Also, a trade representing 10 basis points of a firm's shares is close in size to the 12 basis points used by Keim and Madhavan (1997) as a typical case.

The inputs in Panel A of Table 1 imply additional properties of the benchmark scenario, listed in Panel B. The first of these is the active share of the aggregate active portfolio (AS), which has an implied value of 0.28. This value closely matches the mid-sample trend value from Stambaugh (2014) for the active share of the aggregate portfolio of active equity mutual funds.[7]

The second quantity in Panel B is the ratio of active management's expected trading cost to its AUM, i.e., the value of $\mathrm{E}\left\{\bar{C}/W\right\}$ in equation (32). That quantity's implied value is fairly consistent with the evidence of Edelen, Evans, and Kadlec (2013). Those authors report, for example, that among large equity mutual funds, the average cost per dollar traded is 0.0098, and the average annual turnover is 0.77 of AUM.[8] The product of those values equals 0.0075, comparable in magnitude to the implied value of 0.0066 in Panel B.

The next implied quantity in Panel B is the net alpha earned by investors in active management, i.e., the value of α_{net} in equation (32). Beginning with the seminal study of Jensen (1968), numerous academic studies over the subsequent five decades find negative average net alphas on active equity mutual funds. As summarized by Elton and Gruber (2013), most studies find net alphas that are negative but relatively small, around 0.01 or less in magnitude on an annualized basis. The implied net alpha in Panel B of -0.0054 is well within that range.

Panel B also summarizes the implied dispersion in individual stocks' alphas. Whether this dispersion seems reasonable when compared to, say, returns reported in the anomalies literature, probably depends on how one views the latter, especially with regard to pre- and post-publication evidence. The model's implied difference between average annual alphas in the top and bottom quintiles is 0.031, which is virtually identical to the value of 0.032 reported by McLean and Pontiff (2016) for the average return spread between the top and bottom quintiles when sorting stocks on 97 anomaly variables subsequent to their corresponding publication dates.[9] As compared to that post-publication average, those authors also find that the pre-publication average is about twice as high, but at least part of that higher average could reflect data mining, as the authors discuss. Observe also in Panel B that the average alpha for the bottom decile, -0.022, is more than twice the magnitude of the average for the top decile, 0.009. This asymmetry is consistent with alphas reported in the

[7]As Stambaugh acknowledges, these active-share data were supplied courtesy of Martijn Cremers.

[8]Reported fund turnover, following SEC guidelines, reflects the lesser of a fund's buy volume and sell volume, thereby largely excluding trading activity driven by fund flows. This treatment is consistent with the current model's excluding the costs of flow-driven trading.

[9]That study's Table 1 reports an average post-publication monthly difference of 0.00264, and multiplying by 12 gives 0.032.

anomalies literature. For example, based on a measure combining 11 prominent anomalies, Stambaugh, Yu, and Yuan (2015) find that the negative average benchmark-adjusted return among bottom-quintile stocks is more than twice the magnitude of the positive average return for top-quintile stocks.

An additional implication reported in Panel B is the correlation between the α_i's and an individual active manager's alpha forecasts, i.e., the $\alpha_i^{(j)}$'s from equation (9). This correlation's implied value, equal to 0.24, does not necessarily have an empirical counterpart. It simply provides another measure of an individual manager's imperfect ability to identify alphas, which reflect all managers' collective signals.

4.2. Skill shifts and value added

The quantities listed in Table 1 constitute a broad set of characteristics of the active management industry and asset returns. Within this empirically relevant setting, I explore the effect on a managers' expected value added when some fraction of managers shift their skill. Specifically, I compare an equal-skill equilibrium in which all managers have $\sigma_\eta^{(j)} = \sigma_\eta$ to other equilibria in which a fraction γ of managers have $\sigma_\eta^{(j)} = \sigma_\eta^*$ while the remaining managers have $\sigma_\eta^{(j)} = \sigma_\eta$. The latter value is held constant at 0.06 along with the model's other inputs in Panel A of Table 1.

4.2.1. Managers whose skill shifts

Figure 2 displays implications for a manager whose skill shifts, with $\sigma_\eta^{(j)}$ moving from σ_η to σ_η^*. For a given fraction, γ, of managers whose skill shifts in that manner, each curve traces the equilibria across different values of σ_η^* on the horizontal axis. The vertical axis plots the ratio of the manager's (expected) value added under that equilibrium to the manager's value added under the equal-skill equilibrium in which all managers have $\sigma_\eta^{(j)} = \sigma_\eta$. All curves therefore cross at $\sigma_\eta^* = \sigma_\eta$.

If a manager's skill shifts, what is the impact on her value added? As Figure 2 reveals, the answer depends on γ, the fraction of managers whose skill shifts along with hers. When γ equals 1.00 or 0.90, the manager's value added shifts in the opposite direction as her skill. For example, an increase in skill, corresponding to $\sigma_\eta^* < \sigma_\eta$, is accompanied by a decrease in value added. The reason is as discussed at the outset: First, the increased skill for some managers produces greater price correction in equilibrium and thus smaller magnitudes of stocks' alphas. That effect is greater, the larger is γ. When γ is sufficiently high, the smaller alpha magnitudes outweigh the opposing effect of the manager's increased skill to identify alphas more precisely. As a result, the manager earns lower investment profits. Similarly, a decrease in skill to identify alphas is more than offset by the accompanying lower price correction and larger alpha magnitudes, resulting in higher profits.

The fewer are the managers whose skill shifts, the less that shift impacts prices. With less price impact, and thus smaller effects on stocks' alphas, a manager's increased skill to identify alphas outweighs the opposing effect of smaller alpha magnitudes. As Figure 2 shows, when γ equals 0.25 and 0.50, the manager's value added shifts in the same direction as her skill. For example, an increase in skill ($\sigma_\eta^* < \sigma_\eta$) is accompanied by an increase in value added. That is the case, anyway, for σ_η^* within a sufficiently confined neighborhood of σ_η. As Figure 2 also shows, for sufficiently large increases in skill (sufficiently low σ_η^*), that greater skill can instead produce lower value added, as illustrated when $\gamma = 0.5$. The reason is essentially the same as discussed earlier for the larger γ values. Even when a lower fraction of managers become more skilled, if that increase in skill is large enough, the accompanying increased price correction and lower alpha magnitudes outweigh the skill to identify alphas more precisely.

The middle case in Figure 2, with $\gamma = 0.75$, illustrates an interesting setting in which managers whose skill shifts significantly away from the equal-skill benchmark in either direction produce lower value added. When their skill increases, the resulting smaller magnitudes of alphas outweigh these managers' better identification of alphas. When these managers' skill decreases, their worse identification of alphas outweighs the resulting larger alpha magnitudes.

In a seemingly relevant quantitative setting, the above results illustrate how the effect of a shift in a manager's skill on her value added depends critically on the fraction of managers whose skill shifts along with hers. When that fraction is low, higher skill results in higher value added. When that fraction is large, the opposite effect occurs. For example, suppose a manager's skill increases, with her $\sigma_\eta^{(j)}$ dropping by a third, from 0.06 to 0.04. If that same increase in skill occurs for just 25% of the other managers, then her value added, and that of those managers, increases by over one-half (the y-axis value in Figure 2 is 1.53). If instead that same skill increase occurs for all managers, then the value added of each one drops by over one-fifth (the y-axis value in Figure 2 is 0.78).

4.2.2. Managers whose skill remains constant

Figure 3 parallels Figure 2, except that the vertical axis now reflects the relative shift in value added for a manager whose skill remains at $\sigma_\eta^{(j)} = \sigma_\eta$ while the $\sigma_\eta^{(j)}$ of the other managers shifts to σ_η^*. Recall that managers in the latter category constitute the fraction γ of all managers. For the managers whose skill does not shift, there is no effect opposing the pricing effect of the other managers' skill shift. That shift can work in only one direction from the perspective of the no-shift managers. That is, when the other managers become more skilled, there is more price correction, shrinking the magnitudes of alphas. Those smaller alphas are identified no more precisely than before by the no-change managers, so their profits suffer. Similarly, if the other managers instead shift to lower skill, the resulting larger alpha magnitudes improve the profits of the no-change managers. Therefore all of the curves in Figure 3 slope upward, with a larger γ simply producing a higher slope because the effect on price correction is greater when the skill increase applies to more managers.

4.2.3. All managers: the active management industry

If some managers become more skilled, their value added can increase, as shown in Figure 2. The value added of the remaining managers decreases, as shown in Figure 3. What is the combined effect on expected value added of the active management industry? For the same γ values in those figures, Figure 4 plots the total value added, summed over all managers, again divided by the same quantity in the equal-skill benchmark equilibrium. All of the curves slope upward, meaning that an increase in skill by the fraction γ of managers results in less total value added for the active management industry. This result occurs even for the γ values of 0.25 and 0.50, under which greater value added is produced by the managers whose skill increases (Figure 2). The increase in value added for those managers is more than offset by the drop in value added of the remaining managers due to the greater price correction (Figure 3). When the industry becomes more skilled, its total value added decreases.

4.3. Conditional expectations given signals

Equation (9) assumes that the manager's conditional expectation of a stock's benchmark-adjusted return is proportional to the signal $s_i^{(j)}$. As discussed earlier, this linear conditioning with a zero intercept is a simplification, not a formal implication. The extent to which such an assumption is appropriate depends on the joint distribution of $s_i^{(j)}$ and the equilibrium α_i.

The model, as presented in the previous section, does not make a distributional assumption about $\eta_i^{(j)}$, which is introduced in equation (8) as the deviation of $s_i^{(j)}$ from α_i. Nor does the model assume a distribution for v_i, which is introduced in equation (6) and determines the equilibrium α_i in equation (24). The quantitative specification in this section specifies a distribution for v_i but not for $\eta_i^{(j)}$. Only the standard deviation of the latter, $\sigma_\eta^{(j)}$, is necessary for the implications presented.

Given the Weibull distribution for v_i assumed for this section's quantitative analysis, is there at least an example of a sensible distribution of $\eta_i^{(j)}$ for which equation (9) holds as a reasonable approximation? It seems so, even when not looking beyond the Weibull family. Specifically, suppose the quantity $1 + \eta_i^{(j)}$ has a Weibull distribution with a mean of 1 (because the mean of $\eta_i^{(j)}$ is 0) and a standard deviation of $\sigma_\eta^{(j)}$. The latter is set to $\sigma_\eta = 0.06$ in the benchmark scenario above, and this value, along with the mean, implies a value of k for the Weibull equal to 20.7. Thus, the density for $1 + \eta_i^{(j)}$ is essentially that shown for $k = 20$ in Figure 1. As displayed there, that density is unimodal and moderately left-skewed (skewness = -0.88). I take 10 million draws of v_i from its assumed Weibull distribution, and for each draw I compute the corresponding value of α_i using equation (24) and the benchmark equilibrium value of θ. For each such draw of α_i, I take a corresponding independent draw of $1 + \eta_i^{(j)}$ from its Weibull distribution, subtract 1, and add the result to α_i to obtain $s_i^{(j)}$, following equation (8). I then separate the 10 million draws of $s_i^{(j)}$ into 100 bins defined by their percentiles. Within each bin, I compute the averages of the 100,000 $s_i^{(j)}$'s and their corresponding α_i's.

Figure 5 plots the percentile average of $s_i^{(j)}$ on the horizontal axis versus the corresponding average of α_i on the vertical axis. Note that $\mathrm{E}\{\alpha_i|s_i^{(j)}\} = \mathrm{E}\{r_i|s_i^{(j)}\}$, given that $r_i - \alpha_i$ has zero expectation conditional on the union all managers' signals, by the definition of α_i. Also plotted is the zero-intercept line corresponding to equation (9). The actual averages from the simulation display only slight concavity versus the linear relation. As reported earlier, the Pearson correlation is 0.998, and the overall average α_i is just 2 basis points. The assumption in equation (9) appears reasonable, at least given the ease in finding a distribution of $\eta_i^{(j)}$ for which that equation seemingly provides a good approximation to $\mathrm{E}\{r_i|s_i^{(j)}\}$. Moreover, the moderate negative skewness in that distribution for $\eta_i^{(j)}$ also seems plausible, given that the true alphas in the benchmark quantitative setting are also negatively skewed, as discussed earlier. That is, suppose a fund's stock analysts know that the most severe cases of mispricing are likelier to occur in the direction of overpricing, rather than underpricing. Then it seems reasonable that those analysts' biggest misses (largest absolute $\eta_i^{(j)}$'s) will be misidentifying some stocks as being among those severely overpriced cases.

4.4. Robustness

The model's implications in less realistic quantitative settings are presumably less interesting. Nevertheless, it appears the basic implications about skill's roles are quite robust across alternative specifications. As discussed earlier, the model's inputs for which direct empirical counterparts seem most elusive are the last three in Panel A of Table 1: c, k, and σ_η. For each of these three parameters, I consider two values substantially different from that in Table 1, one lower and one higher, while keeping the other two parameters at their original values. This experiment provides six alternative settings that replace the original benchmark setting. The alternative values of c are 0.01 and 10, versus the original value of 2.0 in Table 1. Instead of the original $k = 0.6$, the alternative k values are 0.2 and 20, which bound the cases displayed in Figure 1. The alternative values of σ_η are 0.01 and 0.20, versus the original value of 0.06.

For each of the six alternative settings, I repeat the previous analyses of skill shifts, wherein a fraction γ of managers has a shift in $\sigma_\eta^{(j)}$ from σ_η to σ_η^*. Figure 6 displays results paralleling those in Figure 2, while Figure 7 parallels Figure 4. (Given these results, those paralleling Figure 3 seem obvious and unnecessary to display.) While the alternative settings generally imply rather different and less realistic values for the quantities in Panel B of Table 1, the results in Figures 6 and 7 tell a similar story to the original analyses. That is, if a skill shift applies to enough managers, each of those managers produces lower value added.

Can becoming more skilled produce higher value added even when all managers become more skilled? In theory, yes, though such a case appears to require substantial departures from the benchmarks in Table 1. For example, Figure 8 shows results for a case with noise-trader demands more highly skewed than in that benchmark setting and with much costlier trading than previously entertained. Specifically, k equals 0.2, its lowest value in Figures 6 and 7, and c equals 40, which is four times its highest value in those figures. The value of σ_η is again 0.06, as in the benchmark setting. For this case, Panel A of Figure 8 displays results for the managers whose skill shifts from from σ_η to σ_η^*, paralleling Figures 2 and 6; Panel B displays the aggregate results for all managers, paralleling Figures 4 and 7. In both panels of Figure 8, observe that all curves slope downward at $\sigma_\eta^* = \sigma_\eta$, even for $\gamma = 1$. This alternative case, however, has some extreme characteristics. Consider the equal-skill setting, for example. With such a high trading-cost parameter, c, the implied ratio of funds' trading costs to AUM is 0.089, more than 13 times the corresponding value of 0.0066 in Panel B of Table 1 that is close to the empirical counterpart. Nevertheless, with the highly skewed noise-trader demands (low k), stocks' equilibrium alphas are extreme enough for managers to produce high value added despite high trading costs. For example, with active management's 80% market share and 1.2% fee rate, as assumed in Table 1, active management's implied annual net alpha is 7.4%, in sharp contrast to the modest negative empirical estimates. With its high value added in this case, active management could lower its net alpha to zero, as in Berk and Green (2004), by charging a much higher 8.6% fee rate while maintaining an 80% market share, or by expanding to a much higher 573% market share while maintaining a 1.2% fee rate, or by effecting some combination of those extremes. Such scenarios seem quite counterfactual for the overall active management industry.

5. Relation to previous studies

A manager with a higher $\sigma_\eta^{2(j)}$ assesses alphas less accurately. The opposing effects in equilibrium of this higher variance have analogs in the model of Grossman and Stiglitz (1980), in which there are opposing effects of higher variance of the signal about the risky asset's payoff. On one hand, higher signal variance in their model reduces the benefit of observing the signal, analogous in the current model to a manager's profits being reduced by identifying stocks' alphas less accurately. At the same time, higher signal variance in their model increases the benefit of observing the signal, because less information is reflected in the risky asset's equilibrium price. The latter effect is analogous in the current model to all managers' profits benefiting from stocks' larger alphas due to less price correction. In their model, the net benefit of observing the signal can either rise or fall with its variance, somewhat analogously to how the effects on a manager's value added of shifting from σ_η to σ_η^* can in principle go in either direction. Grossman and Stiglitz (1980) do not model delegated asset management, but Gârleanu and Pedersen (2018) do so in a model building on the Grossman-Stiglitz framework. Gârleanu and Pedersen show in their model that higher variance of the signal about the risky asset's payoff has an ambiguous effect on aggregate fee revenue for active management. The latter result relates to the analysis in Figure 4, which asks how value added (and thus presumably fee revenue) for the active management industry is affected by higher signal variance for some or all managers.

The current study differs from Gârleanu and Pedersen (2018) in a number of respects. First, I include a quantitative analysis that allows additional insights into the directions of effects that would be more ambiguous if analyzed only qualitatively. Moreover, this analysis examines the effects of skill shifts on not only the value added of the overall industry but also on that of the managers for whom skill shifts and on that for whom it does not.

Second, I assume trading is costly. Empirical evidence indicates trading costs are substantial. As discussed earlier, Edelen, Evans, and Kadlec (2013) estimate that mutual funds' proportional trading costs are roughly comparable to their expense ratios (which support management fees). As Davila and Parlatore (2020) show, depending on the specifics of the setting, price correction can be significantly affected by trading costs in models having both noise traders and information traders. The key friction in the model of Gârleanu and Pedersen (2018), which has costless trading of the risky asset, is the costs investors incur in searching for skilled managers. Manager search can be extensive for institutional investors, as those authors detail. For retail mutual fund investors, the typical search seems often less extensive, given the strong response of fund flows to readily accessible information such as Morningstar ratings (Del Guercio and Tkac (2008)) and *Wall Street Journal* "Category King" designations (Kaniel and Parham (2017)), though mutual funds do spend nontrivially on marketing (Roussanov, Ruan, and Wei (2020)). In any case, the two models differ with respect to the fundamental friction.[10] With that friction being a fund investor's search cost, rather than a fund's convex trading costs, the Gârleanu-Pedersen model also does not imply fund-level decreasing returns to scale, unlike the current model. As noted earlier, empirical evidence supports decreasing returns to scale.

Third, the current model contains many risky assets, instead of one. This feature enables numerous empirical implications that require a cross-section of assets, as discussed earlier. Also, with the cross-section of assets, managers add value by stock selection (choosing weightings across assets) rather than by market timing (trading an economy's single risky asset). As noted by Kacperczyk, van Nieuwerburgh, and Veldkamp (2014), for example, numerous empirical studies find evidence of successful stock selection among investment managers, while evidence of successful market timing is more elusive, though those authors conclude it does occur during recessions.

Stambaugh (2014) also conducts a quantitative calibration for a multi-stock general equilibrium model of the active management industry. That model shares similar specifications of trading costs and noise-trader demands as the current model, but it does not address effects of skill. Rather, it essentially sets $\sigma_\eta^{(j)} = 0$ for all managers.

[10] The Gârleanu-Pedersen model also has an information acquisition cost that is paid once by a manager on behalf of all her investors. Thus, unlike the search cost that each searching investor bears separately, that information cost becomes unimportant with many investors relative to managers, as the authors explain.

6. Conclusions

A fund's expected investment profit depends on individual assets' alphas and the skill with which the fund manager identifies those alphas. When some managers become more skilled, prices get corrected more, shrinking assets' alphas. This effect of greater skill impacts all managers, and it can result in lower expected profit even for the managers whose skill increases, when there are enough of the latter. Even when there are not, and those managers earn higher expected profit, their increased skill can nevertheless result in lower total expected profit for the active management industry. These general equilibrium effects are demonstrated with an empirically relevant specification of the model that accommodates numerous observed characteristics of the active management industry as well as asset returns. Within a given equilibrium and thus a given cross-section of managers, those with greater skill have higher expected investment profit (value added).

In addition to providing the above insights, the model yields a number of empirical implications. First, funds face decreasing returns to scale with respect to their AUM. In addition, funds with higher fee rates are smaller, other things equal, and those funds have portfolio weights that deviate more from benchmark weights. Also, a fund's total trading cost equals its average investment profit, which in turn is essentially its fee revenue if investors earn approximately zero net alpha. As discussed, all of these implications are supported by existing evidence. An empirical implication for future investigation is that funds with higher average total investment profit are those whose active weights have a higher cross-asset correlation with future returns.

Suppose that the active management industry has become more skilled over time, as suggested by the findings of Pástor, Stambaugh, and Taylor (2015). They estimate a proxy for skill (different from that implied here) and observe that its distribution across managers trends upward over the past three decades. In other words, that study's evidence suggests the active management industry has become more skilled. The authors suggest education and technology, for example, could be part of the story. One might even construe the recent trend toward quantitatively managed "smart-beta" products as a self-proclaimed increase in the industry's skill (or at least its "smartness"). The results here show that an increase in overall skill can imply lower expected investment profits and thus lower fee revenue if the former are what must largely support the latter. Fee revenue can decline through a loss of AUM, a drop in fee rates, or both. Recent decades have seen active management lose AUM market share to passive-index competition while the typical fee rate for active management also declined. In the case of equity mutual funds, for example, over the period from 2000 to 2018, active management lost 20% in market share while also reducing its fee rate by 30 basis points.[11]

If greater skill spells less revenue, an upward trend in skill represents a potential challenge for the active management industry in addition to that discussed by Stambaugh (2014), who analyzes the effect of a downward trend in direct equity ownership by individuals, a potential source of noise trading. That is, if not only the presence of noise traders declines, but the mispricing they induce is more skillfully identified, then active management can face a doubly strong headwind in maintaining its large presence in the money management industry. Of course an industry of competing active managers cannot decide to calm that headwind by becoming less skilled. Applying more skill is in each manager's individual interest, because more-skilled managers make more than less-skilled managers, as implied by the model presented here.

A negative effect of skill on industry scale (revenue) has at least an imperfect analogy to any industry that gets more efficient at producing a good or service for which the capacity to consume is relatively constrained. The more efficient the industry becomes in exploiting its productive resources, the less of those resources it needs to employ. A notable example comes from agriculture, where efficiency gains play a big role in that sector's employing a much smaller share of the U.S. labor force than it once did (e.g., Dimitri, Effland, and Conklin (2005)). The capacity for consuming the active management industry's output is constrained in the sense that the industry can go no further than to correct most mispricing. Being more skilled in identifying mispriced assets can enable the industry to accomplish that job with less resources.

While active managers can earn less by becoming better at what they do, due to increased price correction that shrinks stocks' alphas, the greater price correction can have societal benefits. For example, van Binsbergen and Opp (2019) analyze how alphas represent pricing frictions that distort real investment and create efficiency losses. Having prices more accurately reflect underlying fundamentals can allow more efficient allocations of the economy's resources.

[11] Of total equity mutual fund assets, the fraction under active management went from 91% to 71%, while the asset-weighted average expense ratio of active equity funds went from 1.06% to 0.76%. (Investment Company Institute (2019)).

Appendix

Proof of Proposition 1: Define the N-element vectors r, $\tilde{\alpha}^{(j)}$, $\phi^{(j)}$ and ι, whose i-th elements equal r_i, $\tilde{\alpha}_i^{(j)}$ $\phi_i^{(j)}$, and 1, respectively. Also define the $N \times N$ diagonal matrix P whose i-th diagonal element equals p_i. Recall that $d_i^{(j)}$ in equation (4) equals $D_i^{(j)}/p_i$. Substituting into equation (3) and summing over i gives the fund's total trading cost as

$$C^{(j)} = \sum_{i=1}^{N} C_i^{(j)} = \sum_{i=1}^{N} c d_i^{(j)} D_i^{(j)} = c(W^{(j)})^2 \phi^{(j)\prime} P^{-1} \phi^{(j)}. \tag{A1}$$

The fund's benchmark-adjusted return net of trading costs therefore equals

$$r_A^{(j)} = \phi^{(j)\prime} r - C^{(j)}/W^{(j)} = \phi^{(j)\prime} r - cW^{(j)} \phi^{(j)\prime} P^{-1} \phi^{(j)}, \tag{A2}$$

and the manager's objective in equation (16) can be expressed as

$$\max_{\phi^{(j)}} \left(\phi^{(j)\prime} \tilde{\alpha}^{(j)} - cW^{(j)} \phi^{(j)\prime} P^{-1} \phi^{(j)} \right) \quad \text{subject to} \quad \phi^{(j)\prime} \iota = 0. \tag{A3}$$

The corresponding Lagrangian is

$$\mathcal{L} = \phi^{(j)\prime} \tilde{\alpha}^{(j)} - cW^{(j)} \phi^{(j)\prime} P^{-1} \phi^{(j)} - \xi(\iota' \phi^{(j)}). \tag{A4}$$

Differentiating with respect to $\phi^{(j)}$ and solving the first-order condition gives

$$\phi^{(j)} = \frac{1}{2cW^{(j)}} P \left(\tilde{\alpha}^{(j)} - \xi \iota \right). \tag{A5}$$

Because the value-weighted average of the r_i's is zero by construction, $\iota' P r = \iota' P \tilde{\alpha}^{(j)} = 0$, so premultiplying both sides of equation (A5) by ι' and applying $\iota' \phi^{(j)} = 0$ implies $\xi = 0$. Therefore

$$\phi^{(j)} = \frac{1}{2cW^{(j)}} P \tilde{\alpha}^{(j)}, \tag{A6}$$

giving equation (18).

Proof of Proposition 2: I conjecture that p_i takes the form in equation (19), with θ constant across i, and then I verify that this form is consistent with equilibrium. Combining equations (1), (6), and (19) implies that α_i is given by equation (24). Because v_i is assumed to be distributed identically across i, the value of $\sigma_{\alpha,i}^2$ in equation (11) is equal to the same value for all i, σ_α^2 in equation (23), and thus $\lambda_i^{(j)}$ is equal to $\lambda^{(j)}$, given in equation (22). Recall that the aggregate active weight ϕ_i equals $(1/M) \sum_{j=1}^{M} (W^{(j)}/W) \phi_i^{(j)}$, with $W = (1/M) \sum_{j=1}^{M} W^{(j)}$.

Equations (8), (9), (12), and (18) then imply

$$\begin{aligned}
\phi_i &= (1/M)\sum_{j=1}^{M}(W^{(j)}/W)\left(\frac{1}{2cW^{(j)}}p_i\tilde{\alpha}^{(j)}\right) \\
&= \frac{1}{2cMW}p_i\left(\sum_{j=1}^{M}\lambda^{(j)}s_i^{(j)}\right) \\
&= \frac{\lambda}{2cW}p_i\left(\alpha_i+\frac{1}{M}\sum_{j=1}^{M}\frac{\lambda^{(j)}}{\lambda}\eta_i^{(j)}\right) \\
&= \frac{\lambda}{2cW}p_i\alpha_i \\
&= \frac{\mu\lambda}{2cW}(\bar{p}_i-p_i). \qquad \text{(A7)}
\end{aligned}$$

The stock-market wealth of investors is $(1-h)Np_m$, of which MW is invested with active management, and thus $\phi_{S,i}$ in equation (17) equals

$$\begin{aligned}
\phi_{S,i} &= \left(\frac{MW}{(1-h)Np_m}\right)(\phi_{m,i}+\phi_i)+\left(1-\frac{MW}{(1-h)Np_m}\right)\phi_{m,i} \\
&= \phi_{m,i}+\left(\frac{MW}{(1-h)Np_m}\right)\phi_i \\
&= \phi_{m,i}+\left(\frac{MW}{(1-h)Np_m}\right)\frac{\mu\lambda}{2cW}(\bar{p}_i-p_i) \qquad \text{(A8)}
\end{aligned}$$

Recall from equation (5) that $\phi_{H,i}=\hat{p}_i/(Np_m)$, and note $\phi_{m,i}=p_i/(Np_m)$. Substituting these relations and equation (A8) into equation (17) gives

$$\begin{aligned}
\frac{p_i}{Np_m} &= h\phi_{H,i}+(1-h)\phi_{S,i} \\
&= h\frac{\hat{p}_i}{Np_m}+(1-h)\left[\frac{p_i}{Np_m}+\left(\frac{MW}{(1-h)Np_m}\right)\frac{\mu\lambda}{2cW}(\bar{p}_i-p_i)\right]. \qquad \text{(A9)}
\end{aligned}$$

Multiplying through by Np_m and solving for p_i gives

$$p_i=\bar{p}_i+\left(1+\frac{\lambda\mu M}{2hc}\right)^{-1}(\hat{p}_i-\bar{p}_i), \qquad \text{(A10)}$$

which is identical to equation (19) for θ given by equation (20), verifying the initial conjecture.

Proof of Proposition 3: Equation (25) is the same as equation (A7) derived in proving Proposition 2.

Proof of Proposition 4: Recall that active share is $AS=\frac{1}{2}\sum_{i=1}^{N}|\phi_i|$. As N grows large, this quantity is increasingly well approximated as $AS=\frac{N}{2}\mathrm{E}\{|\phi_i|\}$. Combining the latter relation with equations (6) and (19), recalling that v_i is independent of $\bar{p}_i$, gives

$$\begin{aligned}
AS &= \frac{N}{2}\mathrm{E}\{|\phi_i|\} \;=\; \frac{N\mu\lambda}{4cW}\mathrm{E}\{|\bar{p}_i - p_i|\} \;=\; \frac{N\mu\lambda\theta}{4cW}\mathrm{E}\{|\bar{p}_i - \hat{p}_i|\} \\
&= \frac{N\mu\lambda\theta}{4cW}\mathrm{E}\left\{\bar{p}_i\left|\frac{\bar{p}_i - \hat{p}_i}{\bar{p}_i}\right|\right\} \;=\; \frac{N\mu\lambda\theta}{4cW}\mathrm{E}\{\bar{p}_i|1 - v_i|\} \\
&= \frac{N\mu\lambda\theta}{4cW}\mathrm{E}\{\bar{p}_i\}\mathrm{E}\{|v_i - 1|\} \;=\; \frac{N\mu\lambda\theta}{4cW}p_m\mathrm{E}\{|v_i - 1|\} \\
&= \frac{M\mu\lambda\theta}{4c(MW)/(Np_m)}\mathrm{E}\{|v_i - 1|\} \;=\; \frac{M\mu\lambda\theta}{4c\delta(1-h)}\mathrm{E}\{|v_i - 1|\}, \qquad \text{(A11)}
\end{aligned}$$

which is equation (26), with δ as defined in equation (27), recalling $MW = \sum_{j=1}^{M} W^{(j)}$.

Proof of Proposition 5: Define the N-element vectors $s^{(j)}$, α, and $\eta^{(j)}$ whose i-th elements are $s_i^{(j)}$, α_i, and $\eta_i^{(j)}$. Substituting for $\phi^{(j)}$ from equation (A7) into the objective in equation (A3), recalling $V_A^{(j)} = W^{(j)} r_A^{(j)}$, gives the maximized conditional expected value added as

$$\begin{aligned}
\mathrm{E}\{V_A^{(j)}|s_1^{(j)},\ldots,s_N^{(j)}\} &= W^{(j)}\left[\phi^{(j)\prime}\tilde{\alpha}^{(j)} - cW^{(j)}\phi^{(j)\prime}P^{-1}\phi^{(j)}\right] \\
&= W^{(j)}\left[\left(\frac{1}{2cW^{(j)}}P\tilde{\alpha}^{(j)}\right)'\tilde{\alpha}^{(j)} - cW^{(j)}\left(\frac{1}{2cW^{(j)}}P\tilde{\alpha}^{(j)}\right)'P^{-1}\left(\frac{1}{2cW^{(j)}}P\tilde{\alpha}^{(j)}\right)\right] \\
&= \frac{1}{4c}\tilde{\alpha}^{(j)\prime}P\tilde{\alpha}^{(j)} \qquad \text{(A12)} \\
&= \frac{(\lambda^{(j)})^2}{4c}s^{(j)\prime}Ps^{(j)} \\
&= \frac{(\lambda^{(j)})^2}{4c}\left(\alpha + \eta^{(j)}\right)'P\left(\alpha + \eta^{(j)}\right).
\end{aligned}$$

Taking the unconditional expectation and dividing by Np_m to obtain $\Pi^{(j)} = \mathrm{E}\{V_A^{(j)}\}/(Np_m)$ then gives

$$\begin{aligned}
\Pi^{(j)} &= \frac{(\lambda^{(j)})^2}{4cNp_m}\left(\mathrm{E}\{\alpha'P\alpha\} + 2\mathrm{E}\{\eta^{(j)\prime}P\alpha\} + \mathrm{E}\{\eta^{(j)\prime}P\eta^{(j)}\}\right) \\
&= \frac{(\lambda^{(j)})^2}{4c}\left(\frac{1}{N}\sum_{i=1}^{N}\mathrm{E}\left\{\frac{p_i}{p_m}\alpha_i\right\} + 0 + \frac{1}{Np_m}\sum_{i=1}^{N}\mathrm{E}\left\{p_i\eta^{2(j)}\right\}\right) \\
&= \frac{1}{4c}\left(\frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_\eta^{2(j)}}\right)^2\left(\mathrm{E}\left\{\frac{p_i}{p_m}\alpha_i\right\} + \frac{1}{Np_m}\sum_{i=1}^{N}\mathrm{E}\{p_i\}\mathrm{E}\{\eta^{2(j)}\}\right) \\
&= \frac{1}{4c}\left(\frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_\eta^{2(j)}}\right)^2\left(\tilde{\sigma}_\alpha^2 + \frac{1}{Np_m}Np_m\sigma_\eta^{2(j)}\right) \\
&= \frac{\tilde{\sigma}_\alpha^2 + \sigma_\eta^{2(j)}}{4c\left(1 + \sigma_\eta^{2(j)}/\sigma_\alpha^2\right)^2}, \qquad \text{(A13)}
\end{aligned}$$

which is equation (28). From equations (1), (6), and (19), along with the independence of v_i and $\bar{p}_i$,

$$\begin{aligned}
\tilde{\sigma}_\alpha^2 &= \mathrm{E}\left\{\frac{p_i}{p_m}\alpha_i^2\right\} = \mathrm{E}\left\{\frac{p_i}{p_m}\left(\frac{\mu(\bar{p}_i - p_i)}{p_i}\right)^2\right\} = \mu^2\theta^2\mathrm{E}\left\{\frac{1}{p_m}\left(\frac{(\bar{p}_i-\hat{p}_i)^2}{p_i}\right)\right\} \\
&= \mu^2\theta^2\mathrm{E}\left\{\frac{1}{p_m}\left(\frac{(\bar{p}_i-\hat{p}_i)^2}{\bar{p}_i+\theta(\hat{p}_i-\bar{p}_i)}\right)\right\} = \mu^2\theta^2\mathrm{E}\left\{\frac{\bar{p}_i}{p_m}\left(\frac{[(\hat{p}_i-\bar{p}_i)/\bar{p}_i]^2}{1+\theta(\hat{p}_i-\bar{p}_i)/\bar{p}_i)}\right)\right\} \\
&= \mu^2\theta^2\mathrm{E}\left\{\frac{\bar{p}_i}{p_m}\left(\frac{(v_i-1)^2}{1+\theta(v_i-1)}\right)\right\} = \mu^2\theta^2\mathrm{E}\left\{\frac{\bar{p}_i}{p_m}\right\}\mathrm{E}\left\{\frac{(v_i-1)^2}{1+\theta(v_i-1)}\right\} \\
&= (\mu\theta)^2\cdot 1\cdot\mathrm{E}\left\{\frac{(v_i-1)^2}{1+\theta(v_i-1)}\right\}, \qquad \text{(A14)}
\end{aligned}$$

which gives equation (29).

Proof of Proposition 6: Straightforward differentiation reveals that the right-hand side of equation (28) is decreasing in $\sigma_\eta^{2(j)}$ for $\sigma_\eta^{2(j)} > \sigma_\alpha^2 - 2\tilde{\sigma}_\alpha^2$. This inequality holds for all $\sigma_\eta^{2(j)}$ if $2\tilde{\sigma}_\alpha^2 > \sigma_\alpha^2$. From equations (23), (24), and (29),

$$\begin{aligned}
2\tilde{\sigma}_\alpha^2 - \sigma_\alpha^2 &= 2(\mu\theta)^2\mathrm{E}\left\{\frac{(v_i-1)^2}{1+\theta(v_i-1)}\right\} - \mathrm{E}\left\{\left(-\mu\frac{\theta(v_i-1)}{1+\theta(v_i-1)}\right)^2\right\} + \left(\mathrm{E}\left\{-\mu\frac{\theta(v_i-1)}{1+\theta(v_i-1)}\right\}\right)^2 \\
&\geq (\mu\theta)^2\mathrm{E}\left\{\frac{2(v_i-1)^2}{1+\theta(v_i-1)} - \left(\frac{v_i-1}{1+\theta(v_i-1)}\right)^2\right\} \\
&= (\mu\theta)^2\mathrm{E}\left\{\left(\frac{v_i-1}{1+\theta(v_i-1)}\right)^2[1+2\theta(v_i-1)]\right\}. \qquad \text{(A15)}
\end{aligned}$$

A sufficient (not necessary) condition for the above expectation to be positive is that the quantity inside the expectation operator is always positive. The latter occurs when $\theta < 0.5$, because v_i is non-negative.

Proof of Proposition 7: Recall the large-N setting, with sample moments across stocks taken as population expectations. First observe, using equation (18), the properties of the $\eta_i^{(j)}$'s, $\mathrm{E}\{p_i\alpha_i\} = 0$, and $\mathrm{E}\{\phi_i^{(j)}\} = 0$, that the cross-stock covariance of manager j's active weights and returns is given by

$$\begin{aligned}
\mathrm{Cov}\left\{\phi_i^{(j)}, r_i\right\} &= \mathrm{E}\left\{\phi_i^{(j)}, \alpha_i + (r_i - \alpha_i)\right\} = \mathrm{E}\left\{\phi_i^{(j)}, \alpha_i\right\} \\
&= \mathrm{E}\left\{\frac{p_i\tilde{\alpha}_i^{(j)}}{2cW^{(j)}}, \alpha_i\right\} = \frac{\lambda^{(j)}}{2cW^{(j)}}\mathrm{E}\left\{p_i(\alpha_i + \eta_i^{(j)}), \alpha_i\right\} \\
&= \frac{\lambda^{(j)}}{2cW^{(j)}}\left(\mathrm{E}\left\{p_i\alpha_i^2\right\} + \mathrm{E}\left\{p_i\eta_i^{(j)}\alpha_i\right\}\right) = \frac{\lambda^{(j)}p_m}{2cW^{(j)}}\mathrm{E}\left\{\frac{p_i}{p_m}\alpha_i^2\right\} \\
&= \frac{\lambda^{(j)}p_m}{2cW^{(j)}}\tilde{\sigma}_\alpha^2. \qquad \text{(A16)}
\end{aligned}$$

Similarly, the variance of manager j's active weights is given by

$$\begin{aligned}\operatorname{Var}\left\{\phi_i^{(j)}\right\} &= \mathrm{E}\left\{\left(\frac{p_i\tilde{\alpha}_i^{(j)}}{2cW^{(j)}}\right)^2\right\} = \left(\frac{\lambda^{(j)}p_m}{2cW^{(j)}}\right)^2 \mathrm{E}\left\{\left(\frac{p_i}{p_m}\right)^2\left(\alpha_i^2+2\alpha_i\eta_i^{(j)}+\eta_i^{2(j)}\right)\right\} \\ &= \left(\frac{\lambda^{(j)}p_m}{2cW^{(j)}}\right)^2\left(\mathrm{E}\left\{\left(\frac{p_i}{p_m}\right)^2\alpha_i^2\right\}+\mathrm{E}\left\{\left(\frac{p_i}{p_m}\right)^2\right\}\sigma_\eta^{2(j)}\right). \qquad \text{(A17)}\end{aligned}$$

Combining equations (A16) and (A17) gives the correlation between $\phi_i^{(j)}$ and r_i as

$$\operatorname{Corr}\{\phi_i^{(j)},r_i\} = \frac{\tilde{\sigma}_\alpha^2}{\left(\mathrm{E}\left\{\left(\frac{p_i}{p_m}\right)^2\alpha_i^2\right\}+\mathrm{E}\left\{\left(\frac{p_i}{p_m}\right)^2\right\}\sigma_\eta^{2(j)}\right)^{1/2}\left(\operatorname{Var}\{r_i\}\right)^{1/2}}, \qquad \text{(A18)}$$

which is decreasing in $\sigma_\eta^{(j)}$.

Proof of Proposition 8: Combining equations (A1), (A7), and (A12) gives

$$\begin{aligned}C^{(j)} &= c\left(W^{(j)}\right)^2\phi^{(j)\prime}P^{-1}\phi^{(j)} \\ &= c\left(W^{(j)}\right)^2\left(\frac{1}{2cW^{(j)}}P\tilde{\alpha}^{(j)}\right)'P^{-1}\left(\frac{1}{2cW^{(j)}}P\tilde{\alpha}^{(j)}\right) \\ &= \frac{1}{4c}\tilde{\alpha}^{(j)\prime}P\tilde{\alpha}^{(j)} \\ &= \mathrm{E}\{V_A^{(j)}|s_1^{(j)},\ldots,s_N^{(j)}\} \qquad \text{(A19)}\end{aligned}$$

Taking unconditional expectations gives equation (30).

Table 1

Assumed and implied quantities in the benchmark setting

Description	Value
Panel A. Assumed quantities	
Expected annual gross return on the stock market (μ_m)	1.065
Fraction of the stock market owned by noise traders (h)	0.3
Active management's market share (δ)	0.8
Number of active managers (M)	2200
Active management's average annual fee rate (f)	0.012
Proportional round-trip cost of a stock trade ÷ trade's fraction of stock's market cap. (c)	2
Weibull distribution parameter for noise-trader demands (k)	0.6
Standard deviation of noise in active managers' alpha signals (σ_η)	0.06
Panel B. Implied quantities	
Active management's active share (AS)	0.28
Active management's expected annual trading cost as a fraction of AUM (C/W)	0.0066
Active management's annual net alpha (α_{net})	−0.0054
Dispersion in individual stocks' annual alphas (α_i's)	
Standard Deviation	0.015
Average alpha within top quintile	0.009
Average alpha within bottom quintile	−0.022
Difference between top-quintile and bottom-quintile averages	0.031
Correlation of the α_i's with each individual active manager's $\tilde{\alpha}_i^{(j)}$'s	0.24

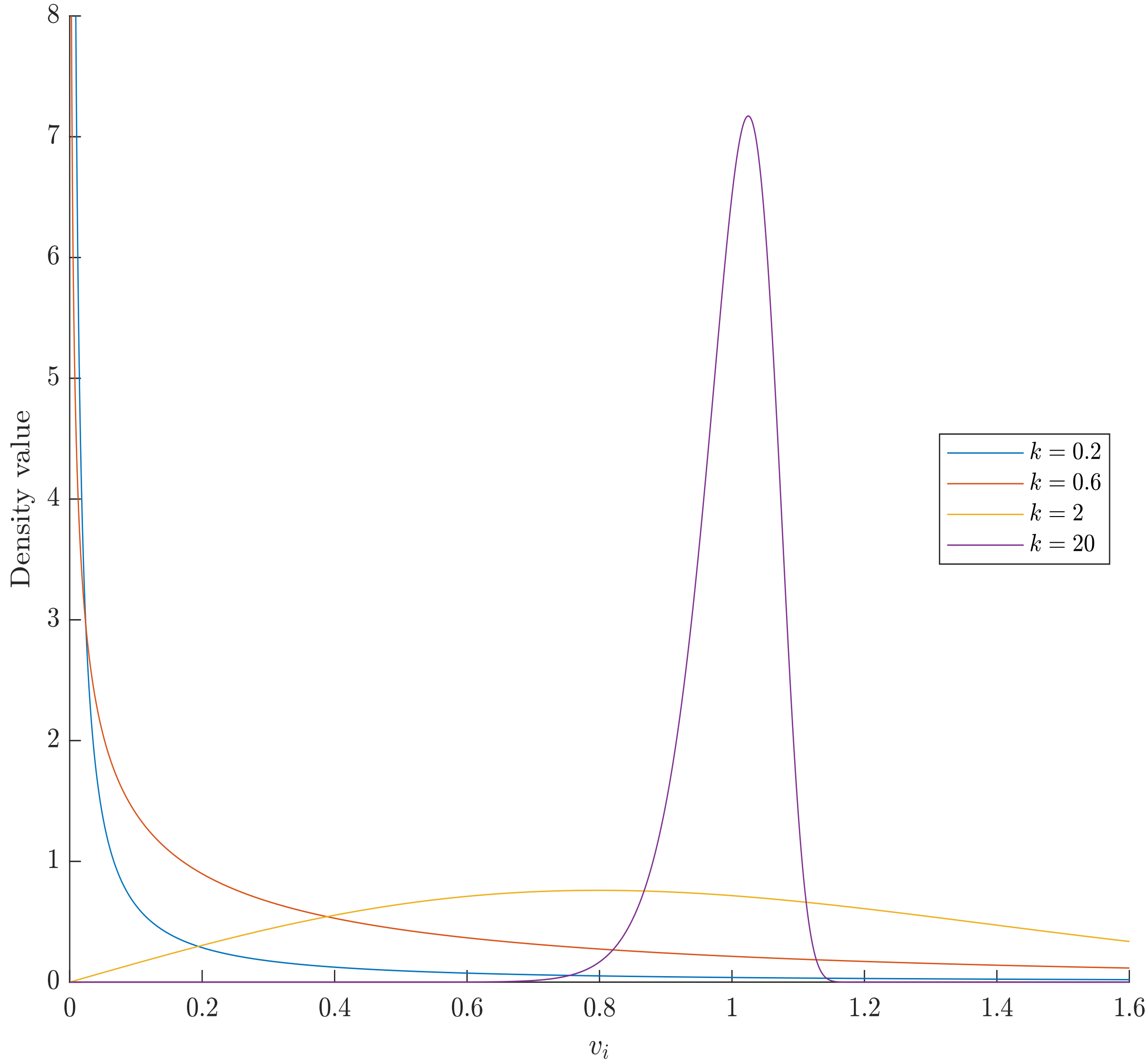


Figure 1. Noise Trading Densities. The figure plots alternative specifications of a Weibull density for v_i, which equals $\hat{p}_i/\bar{p}_i$, where $\hat{p}_i$ is the price of asset i that corresponds to noise-trader demand, and $\bar{p}_i$ is the asset's fair value. All densities have 1.0 as the mean and differ with respect to the shape parameter k.

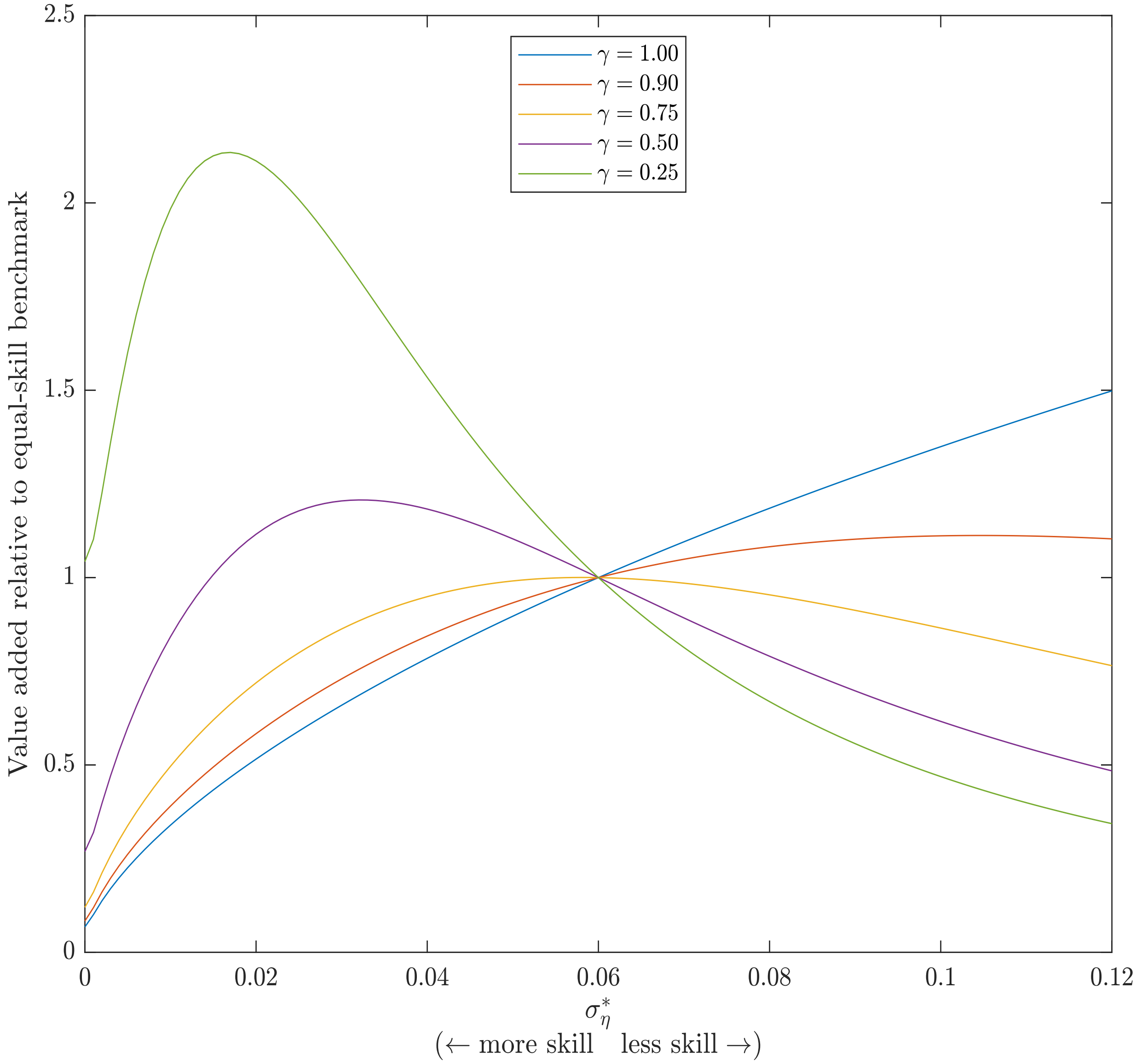


Figure 2. Effects on the value added of the managers whose skill shifts. The value of γ is the fraction of managers whose signal volatility, $\sigma_\eta^{(j)}$, shifts from $\sigma_\eta = 0.06$ to σ_η^*. For the managers whose skill shifts, the vertical axis plots the ratio of their expected value added to that when all managers have $\sigma_\eta^{(j)} = \sigma_\eta$.

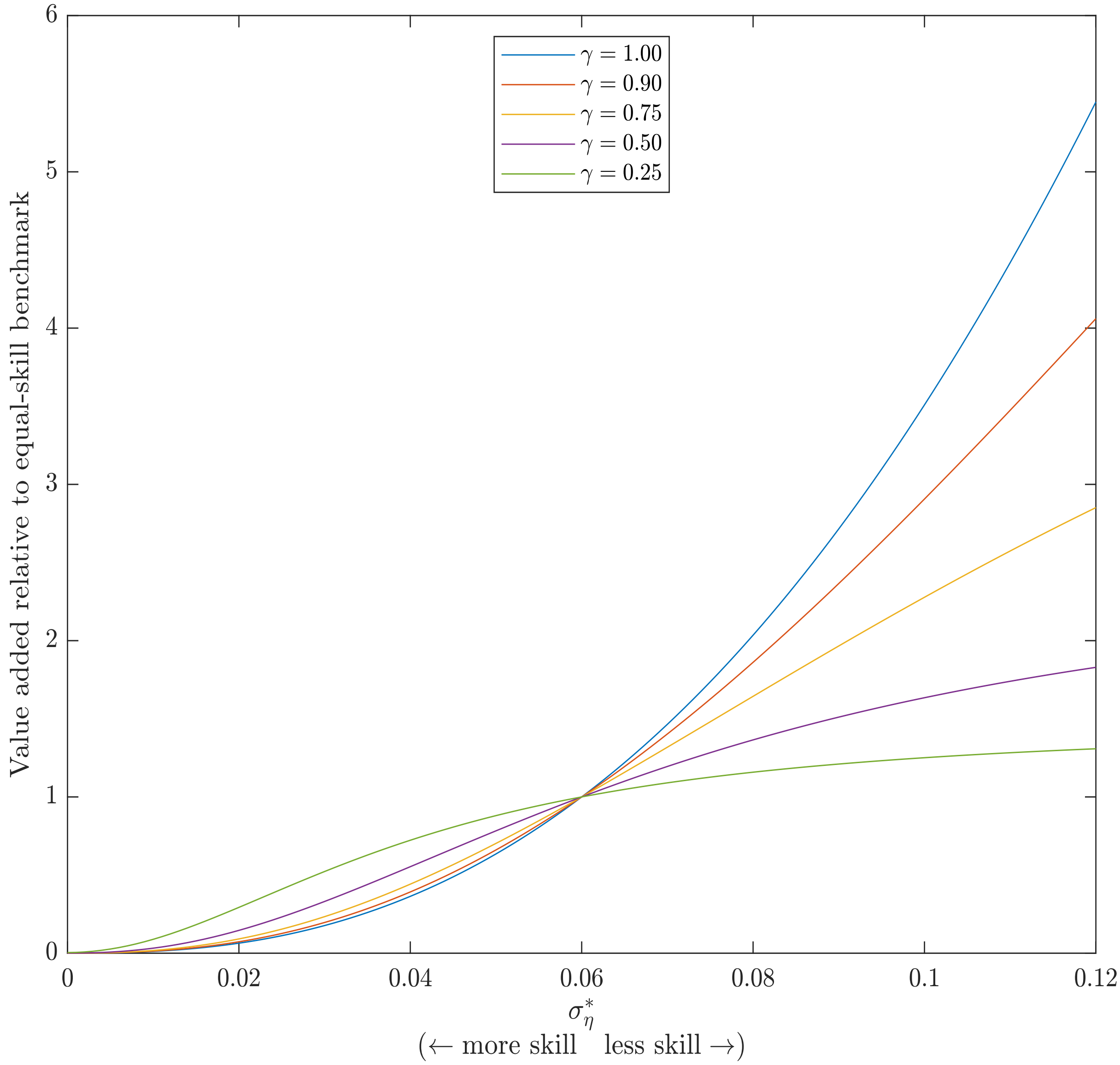


Figure 3. Effects on the value added of the managers whose skill does not shift. The value of γ is the fraction of managers whose signal volatility, $\sigma_\eta^{(j)}$, shifts from $\sigma_\eta = 0.06$ to σ_η^*. For the managers whose skill does not shift, the vertical axis plots the ratio of their expected value added to that when all managers have $\sigma_\eta^{(j)} = \sigma_\eta$.

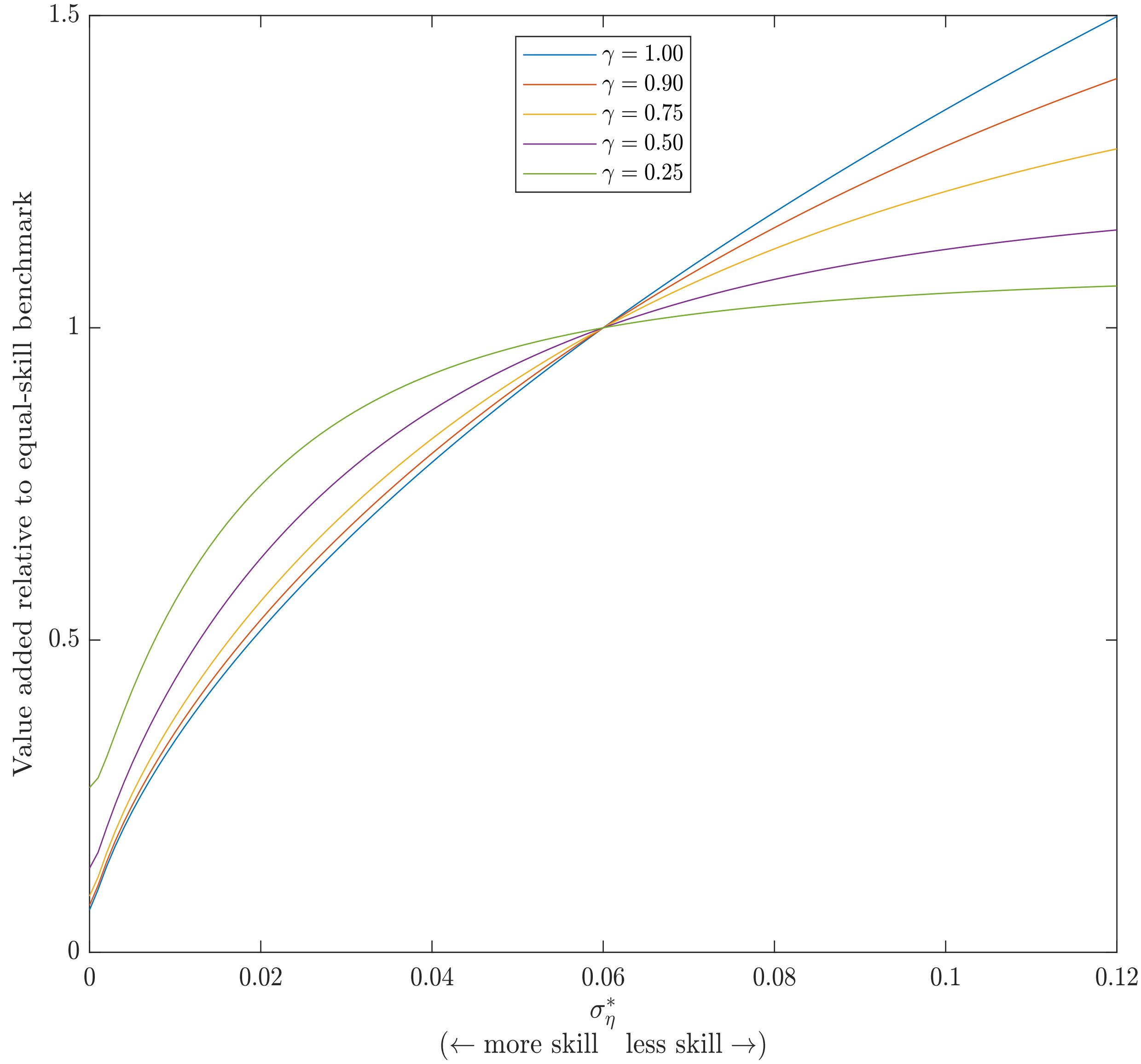


Figure 4. Effects on the total value added of all managers when the skill of some shifts. The value of γ is the fraction of managers whose signal volatility, $\sigma_\eta^{(j)}$, shifts from $\sigma_\eta = 0.06$ to σ_η^*. The vertical axis plots the ratio of the total expected value added for all managers to that when they all have $\sigma_\eta^{(j)} = \sigma_\eta$.

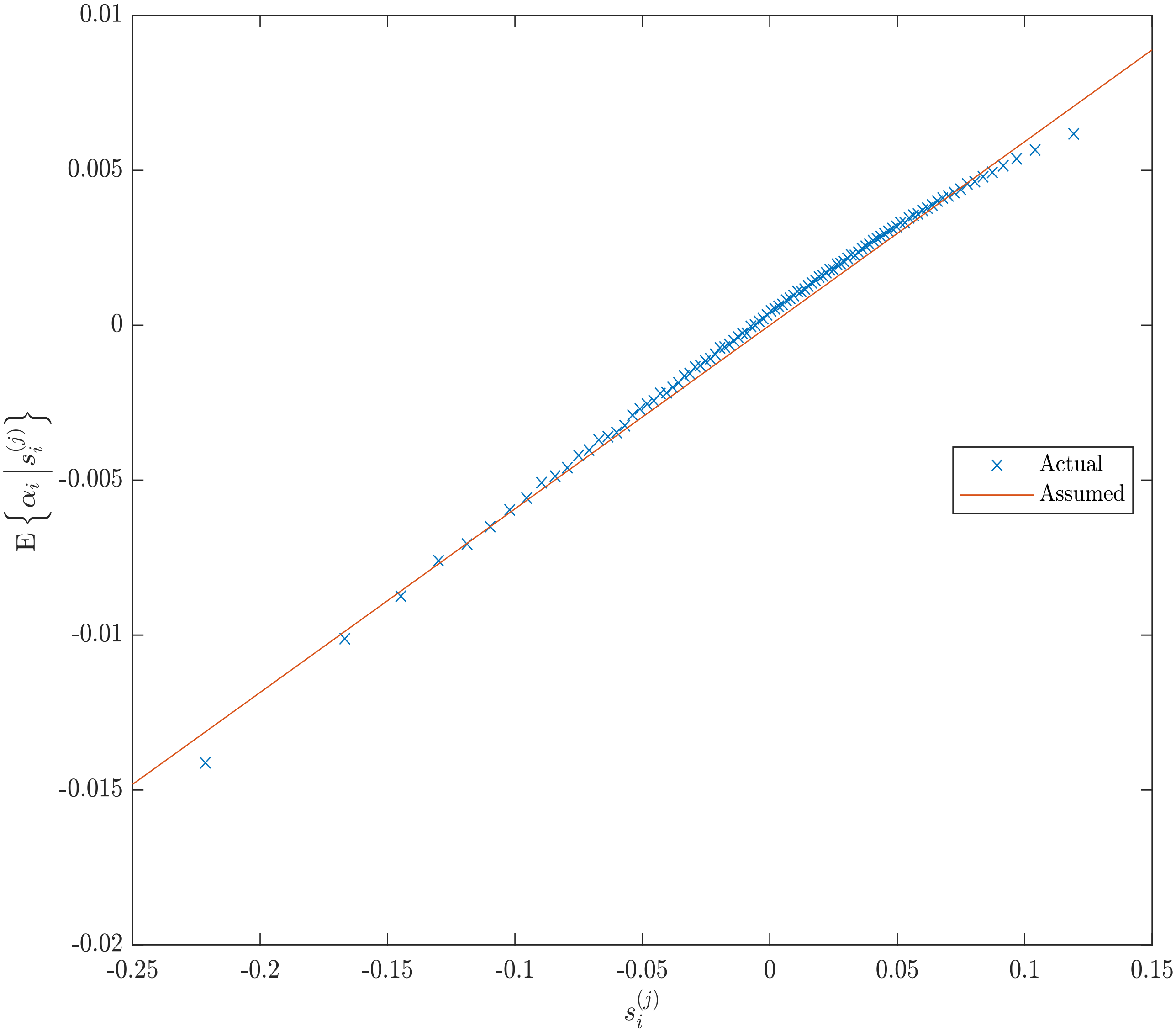


Figure 5. Conditional means of $\boldsymbol{\alpha_i}$ given $\boldsymbol{s_i^{(j)}}$. The individual points plot the model's mean equilibrium α_i within each of the percentiles of $s_i^{(j)}$. The line plots the assumed relation in equation (9).

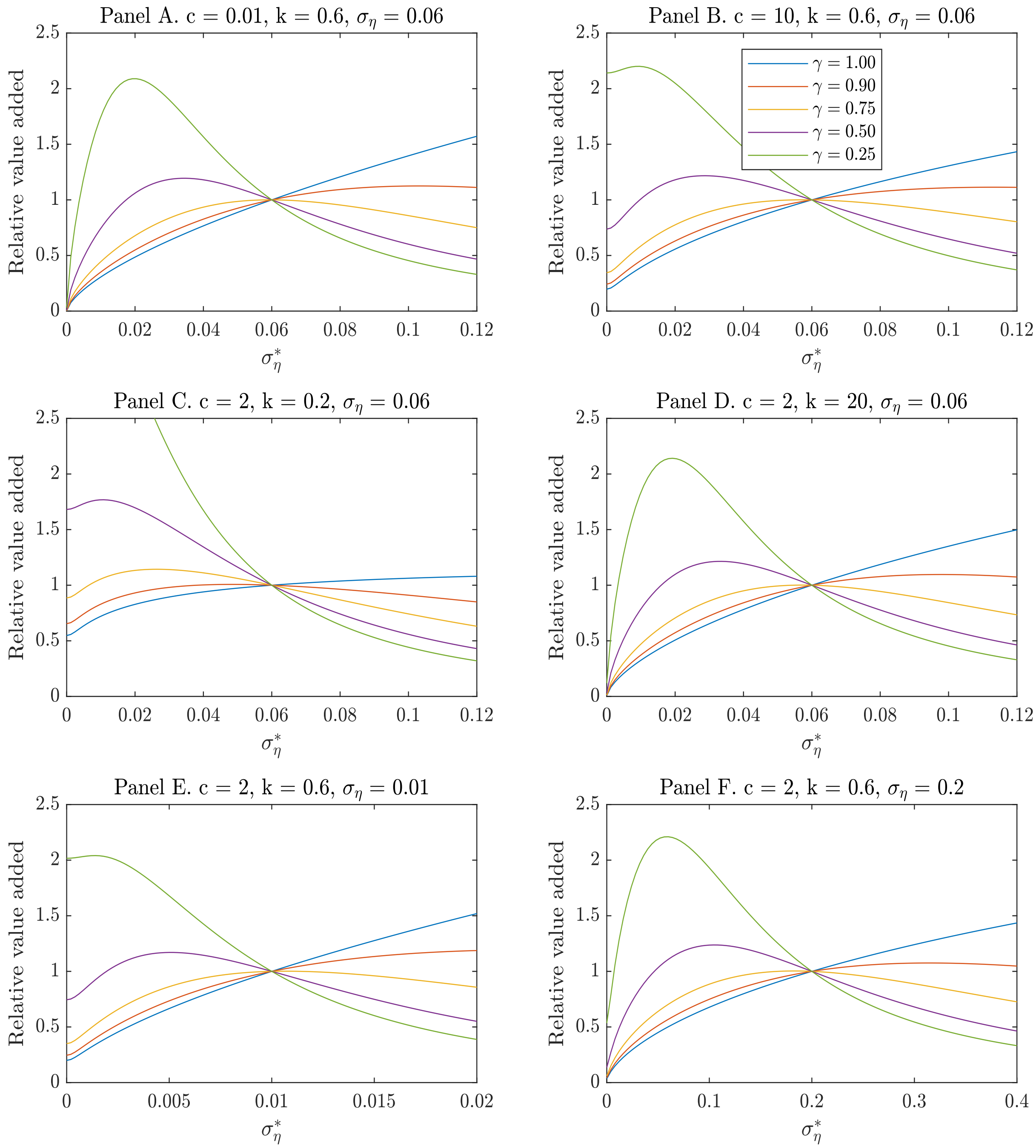


Figure 6. Effects on the value added of the managers whose skill shifts under alternative values of c, k, and σ_η. The value of γ is the fraction of managers whose signal volatility, $\sigma_\eta^{(j)}$, shifts from σ_η to σ^*_η. For the managers whose skill shifts, the vertical axis plots the ratio of their expected value added to that when all managers have $\sigma_\eta^{(j)} = \sigma_\eta$. The value of σ_η is the midpoint of the horizontal axis. The trading-cost parameter is c, and the noise-trader parameter is k.

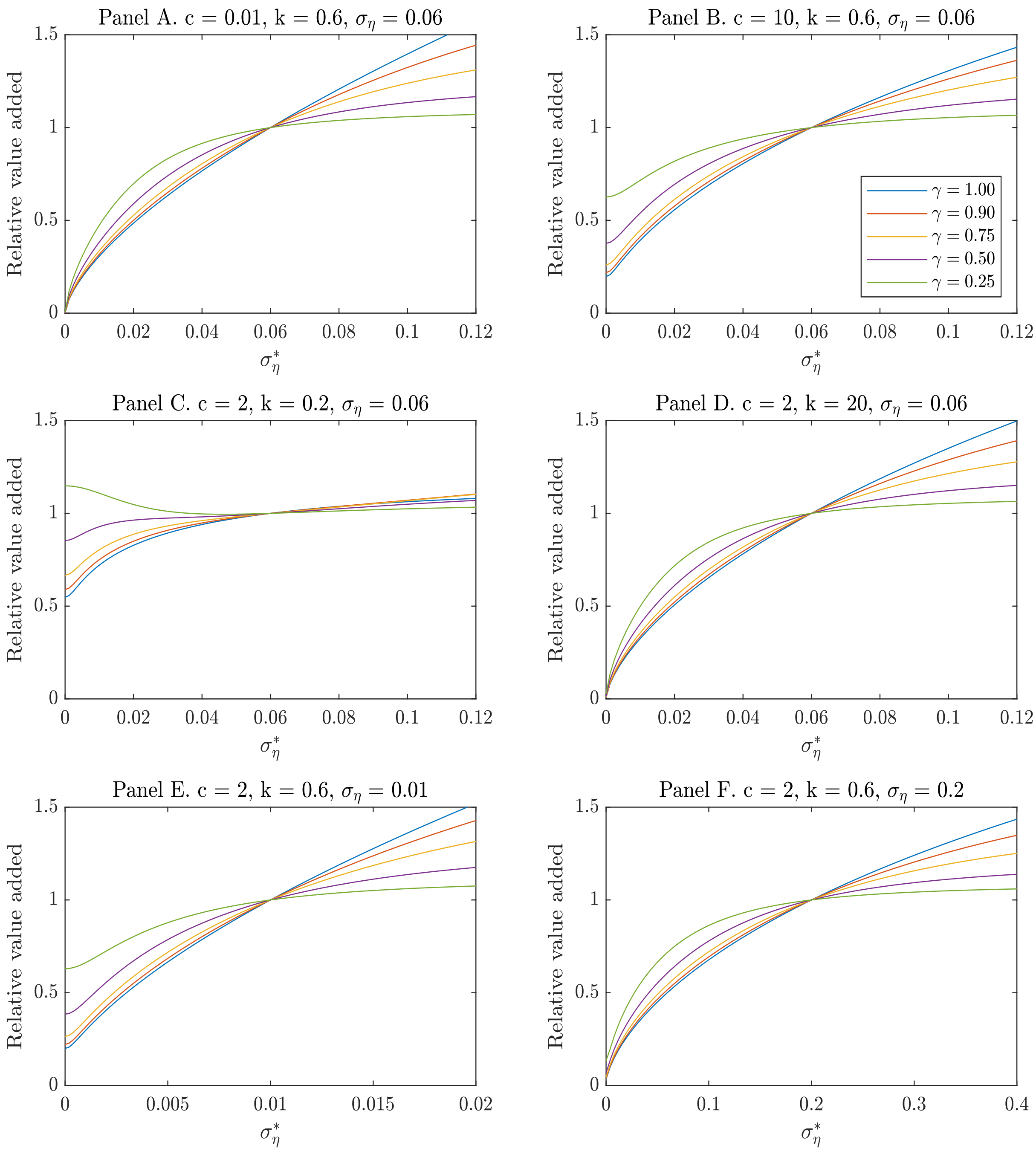
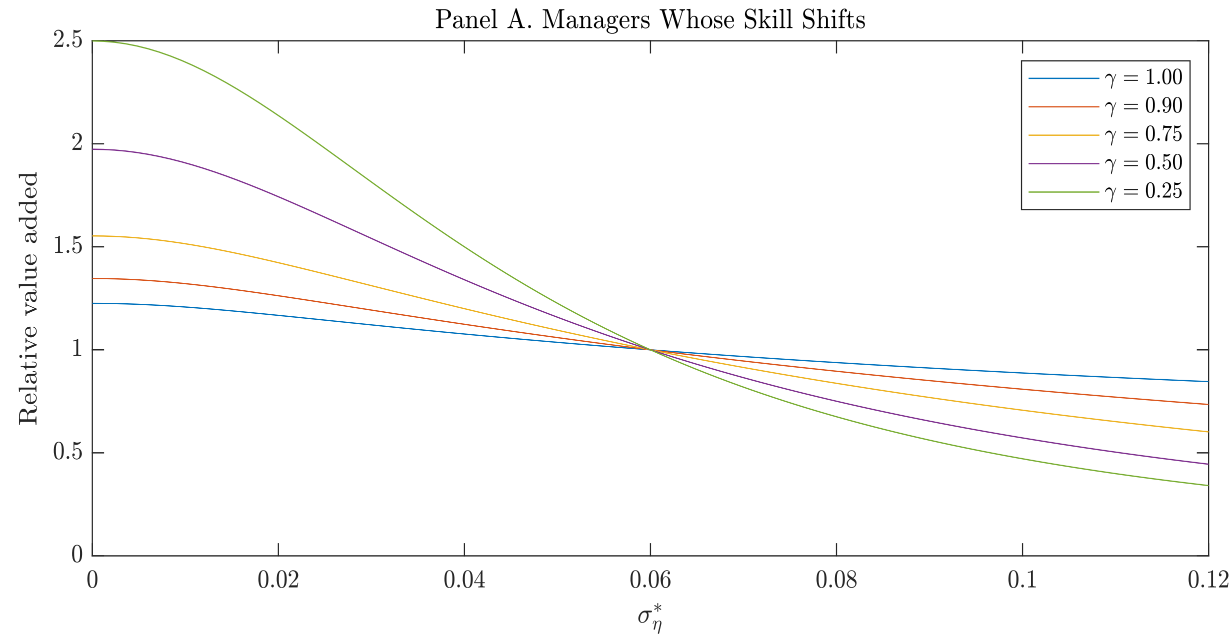


Figure 7. Effects on the total value added of all managers when the skill of some shifts under alternative values of c, k, and σ_η. The value of γ is the fraction of managers whose signal volatility, $\sigma_\eta^{(j)}$, shifts from σ_η to σ_η^*. The vertical axis plots the ratio of the total expected value added for all managers to that when they all have $\sigma_\eta^{(j)} = \sigma_\eta$. The value of σ_η is the midpoint of the horizontal axis. The trading-cost parameter is c, and the noise-trader parameter is k.



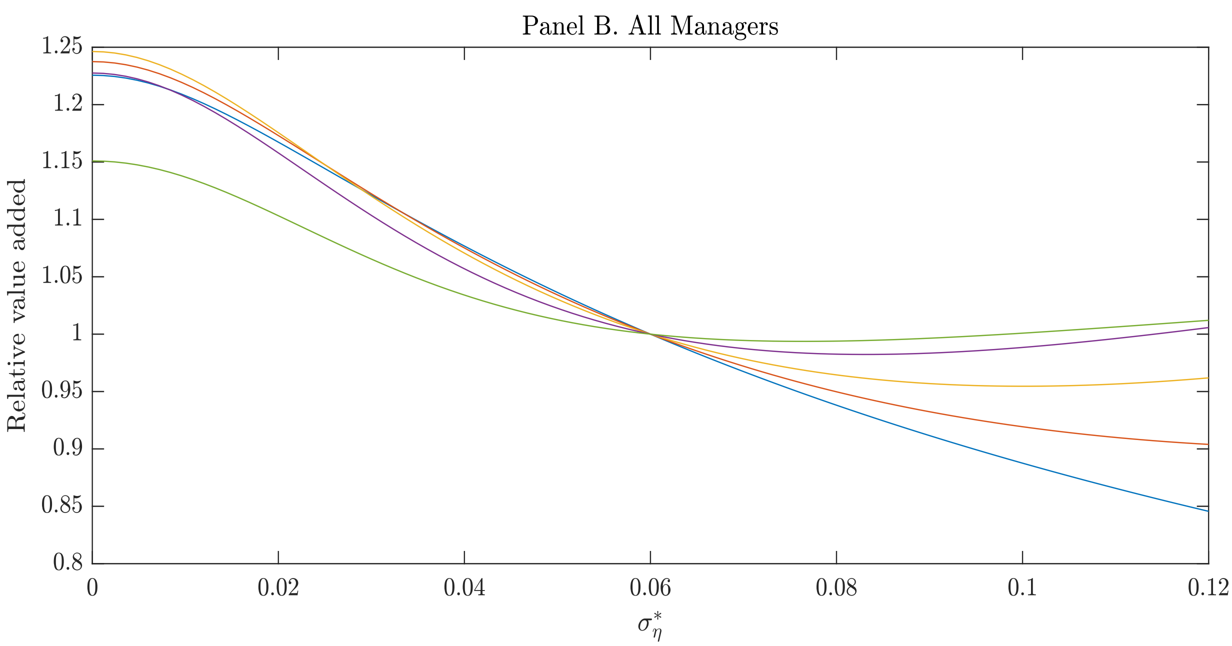


Figure 8. Effects on value added when the skill of some managers shifts under a more extreme setting. The value of γ is the fraction of managers whose signal volatility, $\sigma_\eta^{(j)}$, shifts from $\sigma_\eta = 0.06$ to σ_η^*. For the managers whose skill shifts, the vertical axis in Panel A plots the ratio of their expected value added to that when all managers have $\sigma_\eta^{(j)} = \sigma_\eta$. In Panel B, the vertical axis plots the ratio of the total expected value added for all managers to that when they all have $\sigma_\eta^{(j)} = \sigma_\eta$.

References

Amihud, Yakov, and Haim Mendelson, 1980, Dealership market: Market making with inventory, *Journal of Financial Economics* 8, 31–53.

Banerjee, Snehal, 2011, Learning from prices and the dispersion in beliefs, *Review of Financial Studies* 24, 3025–3068.

Barber, Brad M., and Terrance Odean, 2008, All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors, *Review of Financial Studies* 21, 785–818.

Barber, Brad M., Terrance Odean, and Ning Zhu, 2009, Systematic noise, *Journal of Financial Markets* 12, 547–569.

Berk, Jonathan B., and Richard C. Green, 2004, Mutual fund flows and performance in rational markets, *Journal of Political Economy* 112, 1269–1295.

Berk, Jonathan, and Jules H. van Binsbergen, 2015, Measuring skill in the mutual fund industry, *Journal of Financial Economics* 118, 1–20.

van Binsbergen, Jules H., and Christian C. Opp, 2019, Real anomalies, *Journal of Finance* 74, 1659–1706.

Black, Fischer, 1986, Noise, *Journal of Finance* 41, 529–543.

Blume, Marshall E., Jean Crockett, and Irwin Friend, 1974, Stock ownership in the United States: Characteristics and trends, *Survey of Current Business* 54 (November), 16–40.

Blume, Marshall E., and Irwin Friend, 1975, The asset structure of individual portfolios and some implications for utility functions, *Journal of Finance* 30, 585–603.

Cremers, K.J. Martijn, and Antti Petajisto, 2009, How active is your fund manager? A new measure that predicts performance, *Review of Financial Studies* 22, 3329–3365.

Daniel, Kent, Mark Grinblatt, Sheridan Titman, and Russ Wermers, 1997, Measuring mutual fund performance with characteristic-based benchmarks, *Journal of Finance* 52, 1035–1058.

Dávila, Eduardo, and Cecilia Parlatore, 2020, Trading costs and informational efficiency, Working paper, Yale University and New York University.

Dimitri, Carolyn, Anne Effland, and Neilson Conklin, 2005, The 20th Century transformation of U.S. agriculture and farm policy, Economic Information Bulletin Number 3, United States Department of Agriculture.

Del Guercio, Diane, and Paula A. Tkac, 2008, Star power: The effect of Morningstar ratings on mutual fund flow, *Journal of Financial and Quantitative Analysis* 43, 907–936.

Dorn, Daniel, Gur Huberman, and Paul Sengmueller, 2008, Correlated trading and returns, *Journal of Finance* 58, 885–920.

Edelen, Roger M., Richard Evans, and Gregory B. Kadlec, 2007, Scale effects in mutual fund performance: The role of trading costs, Working paper, Echo Investment Advisors, Boston College, and Virginia Tech.

Edelen, Roger, Richard Evans, and Gregory Kadlec, 2013, Shedding light on "invisible" costs: Trading costs and mutual fund performance, *Financial Analysts Journal* 69 (January/February), 33–44.

Elton, Edwin J., and Martin J. Gruber, 2013, Mutual funds, chapter 15 in Handbook of the Economics of Finance, G.M. Constantinides, M. Harris, and R. Stulz, eds., North Holland, Amsterdam.

Feng, Lei, and Mark S. Seasholes, 2004, Correlated trading and location, *Journal of Finance* 59, 2117–2144.

Garcia, Diego, and Joel M. Vanden, 2009, Information acquisition and mutual funds, *Journal of Economic Theory* 144, 1965–1995.

Gârleanu, Nicolae, and Lasse Heje Pedersen, 2013, Dynamic trading with predictable returns and transaction costs, *Journal of Finance* 58, 2309–2340.

Gârleanu, Nicolae, and Lasse Heje Pedersen, 2018, Efficiently inefficient markets for assets and asset management, *Journal of Finance* 73, 1663–1712.

Glosten, Lawrence R., and Paul R. Milgrom, 1985, Bid, ask, and transaction prices in a specialist market with heterogeneously informed traders, *Journal of Financial Economics* 14, 71–100.

Goetzmann, William N., and Alok Kumar, 2008, Equity portfolio diversification, *Review of Finance* 12, 433–463.

Grinold, Richard, 2006, A dynamic model of investment management, *Journal of Investment Management* 4 (2), 5–22.

Grinblatt, Mark, and Sheridan Titman, 1993, Performance measurement without benchmarks: An examination of mutual fund returns, *Journal of Business* 66, 47–68.

Grossman, Sanford J., and Joseph E. Stiglitz, 1980, On the impossibility of informationally efficient markets, *American Economic Review* 70, 393–408.

Heaton, John, and Deborah J. Lucas, 1996, Evaluating the effects of incomplete markets on risk sharing and asset pricing, *Journal of Political Economy* 104, 443–487.

Investment Company Institute, 2019, *2019 Investment Company Fact Book*.

Jensen, Michael C., 1968, The performance of mutual funds in the period 1945–1964, *Journal of Finance* 23, 389–416.

Johnson, Norman L., and Samuel Kotz, 1970, *Continuous Univariate Distributions – 1*, John Wiley & Sons, New York, NY.

Kacperczyk, Marcin, Stijn van Nieuwerburgh, and Laura Veldkamp, 2014, Time-varying fund manager skill, *Journal of Finance* 69, 1455–1484.

Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng, 2008, Unobserved actions of mutual funds, *Review of Financial Studies* 21, 2379–2416.

Kaniel, Ron, and Robert Parham, 2017, WSJ Category Kings – The impact of media attention on consumer and mutual fund decisions, *Journal of Financial Economics* 123, 337–356.

Keim, Donald B., and Ananth Madhavan, 1997, Transactions costs and investment style: An inter-exchange analysis of institutional equity trades, *Journal of Financial Economics* 46, 265–292.

Kelly, Morgan, 1995, All their eggs in one basket: Portfolio diversification of US households, *Journal of Economic Behavior and Organization* 27, 87–96.

Kyle, Albert S., and Anna A. Obizhaeva, 2016, Market microstructure invariance: Empirical hypotheses, *Econometrica* 84, 1345–1404.

Lease, Ronald C., Wilbur G. Lewellen, and Gary G. Schlarbaum, 1974, The individual investor: Attributes and attitudes, *Journal of Finance* 29, 413–433.

McLean, R. David, and Jeffrey Pontiff, 2016, Does academic research destroy stock return predictability?, *Journal of Finance* 71, 5–31.

Pástor, Ľuboš, and Robert F. Stambaugh, 2012, On the size of the active management industry, *Journal of Political Economy* 120, 740–781.

Pástor, Ľuboš, Robert F. Stambaugh, and Lucian A. Taylor, 2015, Scale and skill in active management, *Journal of Financial Economics* 116, 23–45.

Pástor, Ľuboš, Robert F. Stambaugh, and Lucian A. Taylor, 2020, Fund tradeoffs, *Journal of Financial Economics*, forthcoming.

Polkovnichenko, Valery, 2005, Household portfolio diversification: A case for rank-dependent preferences, *Review of Financial Studies* 18, 1467–1502.

Roussanov, Nikolai, Hongxun Ruan, and Yanhao Wei, 2020, Marketing mutual funds, *Review of Financial Studies*, forthcoming.

Sharpe, William F. 1991, The arithmetic of active management, *Financial Analysts Journal* 47 (January/February), 7–9.

Stambaugh, Robert F., 2014, Presidential address: Investment noise and trends, *Journal of Finance* 69, 1415–1453.

Stambaugh, Robert F., Jianfeng Yu, and Yu Yuan, 2015, Arbitrage asymmetry and the idiosyncratic volatility puzzle, *Journal of Finance* 70, 1903–1948.

Stoll, Hans R., 1979, The supply of dealer services in securities markets, *Journal of Finance* 33, 1133–1151.

Zhu, Min, 2018, Informative fund size, managerial skill, and investor rationality, *Journal of Financial Economics* 130, 114–134.