Market Discipline and Internal Governance in the Mutual Fund Industry

Thomas Dangl
ISK Vienna and Vienna University of Technology

Youchang Wu
University of Vienna

Josef Zechner
University of Vienna

We develop a continuous-time model in which a portfolio manager is hired by a management company. On the basis of observed portfolio returns, all agents update their beliefs about the manager's skills. In response, investors can move capital into or out of the mutual fund, and the management company can fire the manager. Introducing firing rationalizes several empirically documented findings, such as the positive relation between manager tenure and fund size or the increase of portfolio risk before a manager replacement and the following risk decrease. The analysis predicts that the critical performance threshold that triggers firing increases significantly over a manager's tenure and that management replacements are accompanied by capital inflows when a young manager is replaced but may be accompanied by capital outflows when a manager with a long tenure is fired. Our model yields much lower valuation levels for management companies than simple applications of discounted cash flow (DCF) methods and is thus more consistent with empirical observations. (*JEL* G11, G23, G30)

The portfolio management industry has grown substantially over the last few decades, thereby generating increased interest among practitioners, regulators, and academics. The question of efficient governance of delegated portfolio management has attracted special attention. Most theoretical models have addressed this issue using the standard

A previous version of this article was circulated under the title "Mutual Fund Flows and Optimal Manager Replacement." We thank Matthew Spiegel (the editor) and an anonymous referee, who helped us improve the article considerably. We also thank Michael Brennan, Engelbert Dockner, Robert Elliot, Marcin Kacperczyk, Anthony Lynch, Kristian Miltersen, Neal Stoughton, Suresh Sundaresan, Russ Wermers, Stefan Zeisberger, participants of seminars at the University of Vienna, University of Tübingen, City University London, Lancaster University, University of Cologne, Norwegian School of Economics and Business Administration, UCLA, UC Berkeley, Stanford University, HEC Paris, participants of the European Finance Association 2004 meeting in Maastricht, the German Finance Association 2004 meeting in Tübingen, and the American Finance Association 2006 meeting in Boston for valuable suggestions and comments. Youchang Wu and Josef Zechner gratefully acknowledge the financial support by the Austrian Science Fund (FWF) under grant SFB 010 and by the Gutmann Center for Portfolio Management at the University of Vienna. Address correspondence to Josef Zechner, Department of Finance, University of Vienna, Bruennerstrasse 72, A-1210 Vienna, Austria, e-mail: josef.zechner@univie.ac.a.

doi:10.1093/rfs/hhl031 Advance Access publication July 27, 2006

principal-agent paradigm to characterize the optimal contracts that alleviate the agency problem between fund investors and managers.[1] Several empirical studies have investigated the effectiveness of incentive fees[2] and, in a somewhat different vein, the role of funds' boards of directors in controlling agency costs.[3]

This article takes a broader view of the governance mechanisms in the portfolio management industry by simultaneously taking into account internal governance as well as the disciplining effect of the product market. A key characteristic of the mutual fund industry is that fund investors are at the same time consumers. The open-ended structure of mutual funds allows individual investors to "fire" the fund manager by withdrawing their money whenever they feel dissatisfied with the investment management services he provides. This not only disciplines the manager directly, as pointed out by Fama and Jensen (1983), but also gives the management company strong incentives to fire underperforming managers to avoid losing market share. We formalize both product market discipline and internal manager replacement in a fully dynamic framework. Our model shows that the interplay between these two alternative governance mechanisms is the key to understanding many phenomena observed in the mutual fund market.

In our model, the portfolio manager may have some stock-picking ability that generates an abnormal expected rate of return by taking idiosyncratic risks. However, active portfolio management exhibits a diseconomy of scale. Furthermore, the manager's ability to manage a specific fund is unknown to the management company, the fund investors, and the manager himself. All agents in the model start with a common prior distribution of managerial ability and update their beliefs using the observed fund performance. Investors have perfect mobility; they can move money into or out of a fund without any cost. Within such a framework, we address several main questions. First, how much capital is invested into or withdrawn from the fund for a given portfolio performance and how does that fund flow depend on managerial characteristics such as tenure or uncertainty of managerial ability? Second, given these product market forces, what is the optimal manager replacement policy for the fund management company and what are its main determinants? Third, what are typical valuation levels for portfolio management companies and how do they evolve over manager tenure? And fourth, what

[1] See, for example, Stoughton (1993), Heinkel and Stoughton (1994), Admati and Pfleiderer (1997), Das and Sundaram (2002), and Ou-Yang (2003).

[2] See, for example, Elton, Gruber, and Blake (2003).

[3] See, for example, Tufano and Sevick (1997) for the case of open-end funds and Guercio, Dann, and Partch (2003) for the case of closed-end funds.

fund flows and portfolio risk changes are induced by a manager replacement?

We consider both the case in which the precision of the belief about managerial ability remains constant and the case in which the precision of the belief increases over time. For both cases, learning implies a positive and convex relation between unexpected idiosyncratic portfolio returns and fund inflows.[4] In the case in which the precision of the belief increases over time, the fund flow responses are stronger early in a manager's career. By contrast, portfolio returns due to general market movements trigger fund outflows, as documented empirically by Warther (1995) and Fant (1999).

Several of this article's main results focus on the portfolio manager replacement decision. We derive an inverse relationship between the probability of manager replacement and past fund performance. This is in accordance with empirical findings in Khorana (1996) and Chevalier and Ellison (1999a). In our model, manager turnover is more performance sensitive in the early years of a manager's tenure, and managers with longer tenure tend to manage larger funds and have a higher probability to retain their positions. These predictions are confirmed by Chevalier and Ellison (1999a,b) and Fortin, Michelson, and Jordan-Wagner (1999).

The analysis also generates new predictions about manager replacement for which empirical evidence is not yet readily available. Whenever the precision of the belief about managerial ability increases over time, the critical ability level at which firing takes place increases substantially over the manager's tenure. Thus, the management company may find it optimal to fire a manager who is believed to have above-average ability. This is so because the high precision implies that it is highly unlikely that the manager will ever become a "star," even if he is believed to be above average. It is better to hire a new manager with lower expected ability but with more upside potential.

We find that the management company's decision to fire a portfolio manager is accompanied by capital flows and by changes in the risk of the fund portfolio. For most parameter values, a manager replacement is preceded by capital outflows and a portfolio risk increase, then followed by capital inflows and a portfolio risk decrease. However, if a manager with a sufficiently long tenure is fired and the volatility of managerial ability is sufficiently low, then the model predicts the opposite. In such cases, we expect the manager replacement to be followed by capital outflows and an increase in the risk of the fund portfolio. The analysis also

[4] The positive and convex relationship between performance and fund flows has been documented, for example, by Ippolito (1992), Gruber (1996), Chevalier and Ellison (1997), Sirri and Tufano (1998), Bergstresser and Poterba (2002), Boudoukh et al. (2003), and, focusing on pension plans, by Huberman and Sengmüller (2004).

reveals how these responses vary cross-sectionally with parameters such as the uncertainty about managerial talent or the cost of a manager replacement.

In our model, the management company represents a contingent claim on the talent of the portfolio manager. Thus, applying methods from contingent claims pricing, the article provides a consistent framework to analyze the value of the management company. Very little is known about the market values and the value drivers of management companies. In a recent article, Huberman (2004) provided some evidence from stock market listed management companies. The reported valuation levels are generally below 5% of the assets under management, which is much less than traditional DCF valuation methods would imply. By contrast, the real option approach derived in this article produces significantly lower values than the DCF approach, which are thus much more in line with empirically observed values. The numerical analysis also implies a particular time pattern of the value of the management company over the tenure of the portfolio manager.

Our theory is most closely related to Berk and Green (2004). As in their model, we also assume competitive provision of capital by investors, decreasing returns to scale in active portfolio management, and learning about managerial ability via past portfolio returns. We extend their model by explicitly distinguishing between the management company and the portfolio manager. While Berk and Green (2004) focused on a fund with an exogenous shutdown threshold, we analyze a fund management company as a contingent claim on the *posterior* belief about the portfolio manager's ability in an infinite-horizon, continuous-time model. We emphasize the management company's governance role and allow the management company to control the belief process at any time by firing the portfolio manager. This allows us to derive the optimal replacement strategy and empirical fund flow and portfolio risk patterns associated with manager replacement. It also allows us to develop simple valuation expressions for the management company which can be calibrated to empirically observable parameter values.

Lynch and Musto (2003) developed a two-period model to explain the convexity of the flow-performance relation. Their key insight is that underperforming funds will change their strategies, whereas those outperforming will not. Therefore, bad past performance contains less information about future performance than good past performance does. Taking this effect into account, rational investors will be less sensitive to past performance when it is poor. Our model differs from the Lynch and Musto (2003) model in several respects. Lynch and Musto (2003) derived their results in a setting without diseconomies of scale, whereas we allow capital flows to adjust to equate risk-adjusted expected future returns across funds. In this setup, learning about managerial ability

However, a well-diversified fund, such as an index fund, will not be hurt as much by its large size.[7]

1.2 Inference about managerial ability

We assume that neither the fund management company, nor the fund investors, nor the manager himself, can directly observe managerial ability, θ_t. Therefore, the process W_{it} is not observable either. In other words, when agents in our model observe a high (low) market-adjusted NAV return, they cannot be sure whether this is due to good (bad) luck or due to the manager's ability. All the other terms in Equation (1) are assumed to be observable. The agents share a common prior belief: θ_0 is normally distributed with a mean a_0 and variance v_0, and they all use Bayes' rule to update their beliefs as they observe the realized NAV process.

We assume that the true ability, θ_t, is changing randomly over time and can be described by a driftless Wiener process:

$$d\theta_t = \omega dW_{\theta t} \tag{3}$$

where ω is the instantaneous volatility of the true managerial ability, $W_{\theta t}$ denotes a standard Wiener process driving θ_t. $W_{\theta t}$ is assumed to be uncorrelated with the Wiener processes driving the market return and idiosyncratic return, that is, W_{mt} and W_{it}, respectively. Equation (3) is motivated as follows. A manager may improve his investment skill by learning from his experience, and this may lead to an upward drift of θ_t. However, it is well recognized that the business environment changes rapidly over time, implying that old strategies and trading models can be outdated very quickly. Sometimes past experience might even be an obstacle to future success. When these two effects offset each other on average, we end up with a driftless process for managerial ability. Our specification also nests the special case of constant true managerial ability, which corresponds to $\omega = 0$.

To characterize the learning process, we substitute Equation (2) into (1), and move all the directly observable terms to the left-hand side and denote the resulting expression by $d\pi_t$:

[7] Our formulation of the abnormal return can be interpreted in another way. Suppose that the manager divides his assets under management into two parts: one inactive part with systematic risk σ_m and zero idiosyncratic risk and one actively managed part with systematic risk σ_m and idiosyncratic risk σ_ϵ. The weights of these two components are $1 - w_t$ and w_t, respectively. The inactively managed part delivers no abnormal return and does not suffer from any diseconomy of scale. The actively managed part produces an abnormal expected rate of return due to the managerial ability θ_t. But the abnormal return per unit of idiosyncratic risk decreases with the size of the actively managed part. It decreases faster when the idiosyncratic risk is higher, because the portfolio with higher idiosyncratic risk is less liquid. Therefore, the abnormal return of the fund can be written as $\alpha_t = w_t\sigma_\varepsilon(\theta_t - w_t\sigma_\varepsilon\gamma A_t)$. This specification is equivalent to Equation (2) because the idiosyncratic risk of the whole portfolio is $\sigma_{it} = w_t\sigma_\epsilon$.

bounded by ω and we have $\lim_{t\to\infty} v_t = \omega$. Similarly, if $v_0 < \omega$, v_t will increase over time and converge to ω.[9]

Define Z_t as the innovation process of unexpected returns to idiosyncratic risks such that its increment dZ_t is a normalized measure of the deviation of $d\pi_t$ from its *posterior* mean, $a_t\sigma_{it}dt$:

$$dZ_t := \frac{d\pi_t - a_t\sigma_{it}dt}{\sigma_{it}},\ Z_0 = 0.$$

Because Z_t measures the unexpected idiosyncratic return, it represents the signal on which the updating of the belief is based. By construction, Z_t is a standard Wiener process conditional on the common information set of all agents in the model. Unlike the unobservable W_{it} process, the Z_t process is derived from an observable process and is thus observable. Rewriting the dynamics of NAV_t and a_t in terms of dZ_t, we have

$$\frac{dNAV_t}{NAV_t} = [r + \lambda\sigma_m + (a_t\sigma_{it} - \gamma A_t\sigma_{it}^2) - \delta_t - b_t]dt + \sigma_m dW_{mt} + \sigma_{it}dZ_t, \quad (9)$$

$$da_t = v_t dZ_t, \quad (10)$$

where v_t is given by Equation (8). Note that the instantaneous volatility of the *posterior* mean, a_t, is equal to the *posterior* variance, v_t.

1.3 Equilibrium fund size

We assume that after paying a one-time setup cost, the management company charges a proportional fee b_t for its services. The operating cost of managing the fund, including the compensation to the fund manager, is assumed to be a fixed fraction s of the total fee income. Therefore, the instantaneous net profit of the management company is $b_t(1 - s)A_t$. This specification is consistent with a linear sharing rule between the fund manager and the management company.

In the open-end fund market, investors will allocate more capital to funds whose abnormal expected rate of return is higher than the management fee and withdraw money from funds for which the opposite is true. For simplicity, we assume that such fund flows are free of charge, that is, there is perfect capital mobility. Because information is symmetric, investors update their beliefs about managerial ability in the same way as the manager and the management company do, and they monitor the size

[9] If $\omega = 0$, v_t converges to zero, that is, the true ability will be perfectly known as the manager tenure t goes to infinity.

To identify the influence of the belief about managerial ability on fund size, we assume that b_t is constant over time and denote it by b.[12] Clearly, when b_t is fixed at b, σ_{it} and A_t adjust as the belief about managerial ability is updated over time. More specifically, we have the following proposition:

Proposition 1. *The equilibrium size A_t^* of an open-end fund and the optimal level of idiosyncratic risk σ_{it}^* are given by*

$$A_t^* = \begin{cases} \frac{a_t^2}{4b\gamma} & \text{if } a_t > 0 \\ \text{nonexistent} & \text{if } a_t \leq 0 \end{cases} \qquad (12)$$

$$\sigma_{it}^* = \begin{cases} \frac{2b}{a_t} & \text{if } a_t > 0 \\ \text{nonexistent} & \text{if } a_t \leq 0. \end{cases} \qquad (13)$$

Proof. See the discussion preceding the proposition.

Note that the equilibrium fund size A_t^* is a convex function of the *posterior* mean of managerial ability a_t. This explains why the level of idiosyncratic risk σ_{it} is negatively related to the *posterior* mean of managerial ability a_t. Because the marginal return of taking idiosyncratic risk is smaller and decreases faster for larger funds, managers with a higher estimated ability, and thus managing larger funds, find it optimal to take less risk.

Solving for a_t from Equation (13) and substituting into Equation (12), we obtain an equilibrium relation between management fee, fund risk, and fund size:

$$b = \gamma \sigma_{it}^2 A_t. \qquad (14)$$

This equation predicts a linear relation between management fee and the product of the variance of a fund's idiosyncratic returns and fund size, with the coefficient being a measure of the diseconomy of scale. This specification has the advantage that it does not include managerial ability, which is unobservable. Because fee, size, and idiosyncratic risk are all observable, one can directly test this equation. The above specification also suggests a novel way to identify diseconomies of scale. In our setup, the diseconomy of scale does not show up in diminished returns of large funds. However, it can be backed out from the above equation. Because

[12] In practice, management fees are usually very stable. Thus our assumption accords well with empirical evidence.

fund flows can be decomposed into three parts: the expected inflows, the response to unexpected market returns, and the response to unexpected idiosyncratic returns. Both the expected rate of inflows, $v_t^2/4b\gamma A_t - r - \lambda\sigma_m$, and the sensitivity of fund flows to unexpected idiosyncratic returns, $(v_t - b)/\sqrt{b\gamma A_t}$, are positively related to the uncertainty about managerial ability v_t and negatively related to fund size A_t. For reasonable parameterizations, both these terms are positive, implying a positive and convex flow-performance relation.[13] All these results are consistent with empirical findings documented, for example, by Chevalier and Ellison (1997) and Boudoukh et al. (2003).

Equation (18) implies a negative flow response to the fund's returns due to unexpected market movements. Such a relation is documented by Warther (1995) and Fant (1999) at the aggregate level for the mutual fund industry. The different responses of fund flows to market returns and the fund's idiosyncratic returns can be understood in the following way: a fund attracts net inflows if and only if its internal growth rate, namely its NAV return, is less than its equilibrium growth rate determined by changing beliefs. While the positive unexpected market return dW_{mt} has a positive impact on a fund's NAV return, it has no influence on its equilibrium growth rate, because it does not affect the belief about managerial ability. Therefore, its influence on fund size must be offset by corresponding fund outflows.[14] Things are different for the unexpected idiosyncratic return dZ_t. Higher idiosyncratic returns not only increase the asset value but also result in an upward revision of investors' belief about managerial ability, with the latter effect generally dominating the former.

2. Valuation and Manager Replacement Threshold

2.1 The valuation model

In the model developed in Section 1, the state of the world is determined by two variables—the *posterior* mean and the variance of the manager's ability, a_t and v_t, respectively. Because v_t is a deterministic function of manager tenure, the value of the management company, F, can be written as $F = F(a,t)$.

Given the relation between fund size and the belief about managerial ability, it may become optimal for the management company to replace

[13] See Table 2 for our base-case parameterization.

[14] In our model, the investment opportunities for active portfolio management are independent of the size of the entire market. In practice, these opportunities are likely to increase with the size of the market. However, as long as the diseconomies of scale of active fund management are not completely offset by new investment opportunities arising when the market expands, the documented negative relation between market driven fund returns and fund flows will still obtain.