Available online at www.sciencedirect.com

ScienceDirect

Journal of Economic Theory 183 (2019) 907–951

JOURNAL OF Economic Theory

www.elsevier.com/locate/jet

Capital immobility and the reach for yield

Alan Moreira [1]

University of Rochester, United States of America

Received 26 July 2017; final version received 20 June 2019; accepted 23 July 2019
Available online 31 July 2019

Abstract

I build a model in which financial intermediation slows down capital flows. Investors optimally learn from intermediary performance to allocate capital toward profitable intermediaries. Intermediaries reach for yield—i.e., they invest in high-tail-risk assets—in an attempt to drive flows and reduce liquidation risk. Intermediaries with strong opportunities face a trade-off between choosing a portfolio that maximizes profitability, and choosing one that maximizes the speed at which capital flows. In equilibrium, reaching for yield is stronger among intermediaries with weak opportunities, resulting in a reduction in the informativeness of performance; investors thus take longer to learn, and capital flows become less responsive to performance. Capital becomes slow-moving because the reach for yield dampens learning. The model predicts capital immobility to be stronger when tail risk is high; when tail risk is under priced; and in asset classes with large cross-sectional variation in tail-risk exposures.

JEL classification: D82; G14; G23

Keywords: Slow-moving capital; Limit-to-arbitrage; Financial intermediation; Reputation concerns

1. Introduction

Despite the large and increasing share of wealth managed by financial intermediaries, a growing body of work documents that financial capital moves slowly (Duffie, 2010; Pedersen et al.,

E-mail address: alan.moreira@simon.rochester.edu.

[1] I thank Nick Barberis, Dmitry Orlov, Douglas Diamond, Michael Gofman, Gary Gorton, Zhiguo He, Ron Kaniel, Juhani Linnainmaa, Alexi Savov, Laura Veldkamp (the editor), Hongjun Yan, and three anonymous referees for the invaluable feedback.

https://doi.org/10.1016/j.jet.2019.07.010

2007). At the same time, there is an emerging consensus that reaching for yield is widespread in the financial sector (Rajan, 2005, 2008, 2012; Stein, 2013). Financial intermediaries seem to over invest in high-yield high-tail-risk assets.[2] These facts have important aggregate implications on their own, but they are also deeply intertwined. How can investors know where to allocate their capital, if intermediaries can manufacture “alpha” by loading on low-probability tail risks?

In this paper, I build a dynamic model of financial intermediation to study this connection. The model is centered on the interaction between intermediaries and investors’ capital allocation decisions. The key result of this model is that the reach for yield by intermediaries leads to capital immobility. Capital is endogenously slow-moving as a result of yield chasing by some intermediaries. The basic mechanism works as follows: financial intermediaries have strong incentives to improve short-term performance by loading on tail risks. Better performance improves their track record, attracts capital, and reduces the risk of liquidation. Incentives are particularly strong for intermediaries without good investment opportunities. These “opportunistic” intermediaries reach for yield more, dampening the performance advantage of “skilled” intermediaries. Intermediary performance becomes less informative about the underlying quality of the intermediary. Investors take longer to learn, and capital misallocation persists for longer.

Thus, capital immobility is an endogenous result of investors’ optimal response to the incentives intermediaries face. The key assumption is that investors cannot directly measure tail risks in the intermediary portfolio. As a result, loading on tail risks is particularly attractive to an intermediary. It boosts short-term performance without appearing in easy-to-measure forms of risk (like return volatility). The reach for yield delays the speed of learning to the extent that it is stronger among opportunistic intermediaries, which always happen in equilibrium. Intuitively, if all intermediaries were equally aggressive in their reach for yield, performance in a tail event would be completely uninformative. In this case, a more aggressive investment on high-tail-risk assets would enable the opportunistic intermediary to attract flows in the short term, without impacting flows once a tail event hits. Thus, in equilibrium, the opportunistic intermediary must reach for yield more aggressively, resulting in less learning in the short term and more learning during a tail event. Because the opportunistic intermediary optimally balances out these competing forces, the result is that the reach for yield by the opportunistic intermediary reduces the speed of learning and the speed of capital flows.

Reaching for yield by the skilled intermediary has the opposite effect. It increases the speed at which investors learn because it makes it harder for the opportunist to keep up with the performance of the skilled type. However, it generates inefficient variation in investment flows across assets, reducing the expected returns earned by fund investors. Intuitively, as the skilled intermediary concentrates investments in high-tail-risk assets, the portfolio becomes poorly diversified, resulting in a fall in the portfolio’s Sharpe ratio and expected returns. Assets that are heavily exposed to tail risks are quickly inundated with capital as the intermediary attracts new capital flows. Assets that pay well during a tail event experience the opposite phenomenon. As investors pour capital into the intermediary, capital flows into these assets at a very low rate. Thus, when

[2] For example, Rajan (2005) argues: “A number of insurance companies and pension funds have entered in the credit derivative market to sell guarantees against a company defaulting. Essentially, these investment managers collect premia in ordinary times from people buying the guarantees. With very small probability, however, the company will default, forcing the guarantor to pay out a large amount. The investment managers are, thus, taking on tail risks, which produce positive returns most of the time as a compensation for a rare very negative return. These strategies have the appearance of producing very high alphas, so managers have incentives to load on them.”

the skilled intermediary reaches for yield more aggressively, the future allocation of capital improves faster, but the present allocation of capital gets worse.

An example is useful to illustrate how the decisions of the skilled intermediary shape the allocation of capital. Suppose the maximum alpha of the skilled type is 5% and it is achieved with a zero tail exposure. Suppose further that any intermediary can manufacture 5% alpha by loading on tail risks. If only the opportunist reaches for yield, short-term performance differences across types disappear. As a result, learning stalls and capital ceases to move in the short term. Given the allocation of capital to each intermediary, the allocation of capital is the most efficient as it achieves the highest expected return. However, from a dynamic perspective, the allocation of capital is the worst as capital ceases to move to the intermediary that allocates capital better. Now consider the case in which the skilled type also reaches for yield. The fund alpha declines to 4% as her portfolio deviates from the maximum expected return portfolio, but the short-term performance increases to 7% as the portfolio becomes more concentrated on high-tail-risk assets that over perform in the short term. In this case, short-term performance is again informative and capital more mobile. Capital flows faster to the good capital allocator, even though expected returns are lower and the allocation of capital is worse. This trade-off between the present and the future allocation of capital is a key insight from the model.

A novel insight that emerges from the model is the feedback loop between the reach-for-yield behavior of both investors and intermediaries. The closer an intermediary is to liquidation, the stronger the incentives to reach for yield. The increase in incentives to reach for yield is particularly strong among good intermediaries who have the most to lose from liquidation. The relative change in incentives drives investors to expect larger short-term performance differences across intermediaries. Thus, after bad performance, investors naturally respond more aggressively to short-term performance as short-term performance becomes more informative. This response further amplifies the incentives to reach for yield. This reach-for-yield spiral implies that flow-performance sensitivity is a nonlinear function of past performance.

The model makes several predictions that apply to investment vehicles that raise demandable equity capital from arms-length investors—such as mutual funds, hedge funds, and some private equity funds—and that have flexibility in their investment mandate to choose their assets, i.e., the theory applies to active managers. The theory does not require that these intermediaries have the ability to directly make clean tail-risk bets using options, but simply requires that there is cross-sectional variation in asset tail risk that can be identified by the expert but not by her investors.[3]

The first set of predictions relates capital immobility to fund characteristics. Funds with more flexible investment mandates, or with mandates to invest in assets with relatively greater cross-sectional variation in tail risk, exhibit a weaker and more concave relationship between flows and performance. Intuitively, an opportunity set with greater variation in tail risk provides the intermediary with more freedom to invest in high-tail-risk assets without changing observable measures of portfolio risk. There is evidence that these predictions hold in the data for asset classes in which the manager has more flexibility, such as private equity (Kaplan and Schoar, 2005) and hedge funds (Goetzmann et al., 2003), as well as for asset classes with a relatively large cross-sectional variation in tail risk, such as corporate bonds (Goldstein et al., 2015).

[3] See, for example, work by Kelly and Jiang (2014) that documents large cross-sectional variation in tail-risk exposures across U.S. equities.

The model provides a novel explanation for the evidence in Kacperczyk and Schnabl (2012), who documented a large increase in reach for yield across money market fund managers and investors, and who attributed this increase to a lack of market discipline.[4] In the model, investors' understanding of fund manager incentives amplifies the reach-for-yield behavior of both investors and managers. It is the market discipline imposed by investors—i.e., the threat of liquidation—that drives the rampant reach-for-yield behavior.

Perhaps even more strikingly, the model is consistent with the very persistent overpricing of senior tranches of collateralized debt obligations (CDOs) and the concomitant underpricing of junior tranches documented by Coval et al. (2009). Intuitively, the model implies capital flows faster toward the most senior tranches that have relatively higher tail exposure. This results in under investment in the junior tranches and over investment in the senior tranches.

The model has several implications for financial stability. First, tail risks tend to concentrate in the portfolios of financial intermediaries. Second, it suggests that capital reallocation is particularly slow when tail events are more likely. Third, tail risks are more likely to build up in relatively low-risk asset classes, where volatility is a particular poor proxy for tail risks.[5] Fourth, the model predicts that reaching for yield is stronger in low-interest-rate environments, which is consistent with recent empirical evidence (Choi and Kronlund, 2014) and the view of leading policymakers (Rajan, 2005, 2012; Stein, 2013).

The remainder of the paper is organized as follows. After a brief discussion of the literature, Section 2 presents the model setup and characterizes the model solution. I study two economies: a benchmark economy where tail risk is readily observable by investors, and an economy where tail risk cannot be directly measured. In Section 3, a numerical calibration is used to illustrate the implications of the model.

Literature review. This paper relates to a growing body of literature that focuses on implicit incentives that are induced by investor behavior. Chevalier and Ellison (1997); Basak et al. (2007); Chapman et al. (2009); and Basak and Makarov (2014) studied implications for manager portfolio choice, while Brennan and Li (1993), Shleifer and Vishny (1997), Vayanos (2004), Cuoco and Kaniel (2011), Basak and Pavlova (2013), and Kaniel and Kondor (2013) studied the implications of these implicit incentives for asset pricing. These authors took the behavior of investors as given and studied the implications for portfolio choice and equilibrium pricing.

A second strand of literature relevant to this paper studies learning in money management. Berk and Green (2004) showed that the behavior of fund flows and lack of persistence in fund performance could be explained by investors' use of fund performance to learn about their managers. Pastor and Stambaugh (2010) relied on a similar idea to explain the dynamics of the size of the money-management industry. Berk and Stanton (2007) built on these ideas to explain the closed-end fund discount, while Dangl et al. (2008) studied the effect of learning on the optimal replacement of a manager. More broadly, Kozlowski et al. (2016) showed that tail risk can slow down learning substantially. They used this connection to explain the slow recovery after the 2008 financial crises. My work is very different, in that it highlights how tail risks and intermediaries interact in a way that endogenously slows down learning.

This paper connects these two sides of the literature, and studies an environment where learning is endogenous to the manager's trading behavior. The previous work that emphasized the

[4] For example, investors either believed in an implicit government guarantee or neglected the magnitude of the tail risks.

[5] Extreme examples of this disconnect between volatility risk and tail risk are the money market funds studied in Kacperczyk and Schnabl (2012) and the different CDO tranches analyzed by Coval et al. (2009).

dynamics of investor learning has mostly abstracted from the interaction between portfolio choice and learning by either simplifying the investment opportunity set or arguing that the investment opportunity set was "sufficiently non-stationary," which made this type of endogenous response by fund investors infeasible (Shleifer and Vishny, 1997). This paper contributes to the agency literature by showing that learning and manager incentives interact in powerful ways. A novel feedback loop between the reach-for-yield behavior of investors and managers emerges, producing amplification and time variation of reach-for-yield incentives.

My paper is more closely related to the literature that connects the agency and learning views, which is founded on the signal-jamming framework of Holmstrom (1999). Dasgupta and Prat (2008) and Dasgupta et al. (2011) studied the effects of fund managers' reputation concerns on asset pricing in a model where prices were determined by a market maker. Vayanos and Woolley (2013) showed that learning about manager efficiency had the potential to explain the momentum effect. Acharya et al. (2013) also developed a model in which reputation concerns slowed down the identification of good managers. The authors assumed that learning had to start again every time the manager switched to a new project, creating an incentive for managers to switch projects inefficiently to mitigate their reputation risk. This paper focuses instead on the choice of the payoff distribution and applies more directly to financial intermediaries. Makarov and Platin (2015) studied optimal contracting in the case of symmetric information between investors and managers. My paper emphasizes the role of asymmetric information and the link with the speed of capital flows, but takes the contractual environment as given.

The papers most related to my work are Guerrieri and Kondor (2012), Malliaris and Yan (2015), and Di Maggio (2014). These papers studied the impact of reputation concerns on the willingness of a fund manager to invest in strategies that paid well with a low probability. Guerrieri and Kondor (2012) emphasized excess volatility, i.e., how changes in reputation concerns can drive risk premia. Malliaris and Yan (2015) showed that reputation leads managers to avoid bets that pay poorly most of the time. Importantly, none of these showed how reach for yield lead to capital immobility, i.e., a reduction in the speed of capital reallocation to good capital allocators. My paper is the first to show a fundamental trade-off between static and dynamic capital allocation. Specifically, when skilled intermediaries over-allocate to high-yield assets, they reduce the quality of capital allocation in the short term, but they increase the speed at which capital allocation improves.

2. Model

Let's consider that time is continuous $t \in [0, \infty)$ and the economy is populated by a large mass of investors (denoted by I) and one financial expert (denoted by E).[6] Both agents are risk-neutral and discount the future at the risk-free rate of interest ρ.

Investors supply capital to the financial expert. The expert has no wealth, earns zero if unable to attract capital, and invests in a risk-free technology, whose net return is ρdt, and in n risky technologies, whose net returns in excess of the risk-free rate are given by

$$dR_t^\theta = \mu_\theta dt + \sigma dB_t - \left[(\sqrt{\phi}Z_t + \kappa)dJ_t - \lambda\kappa dt\right], \tag{1}$$

where θ denotes the expert type, discussed in detail below, and μ_θ is the vector of expected returns of the risky technologies; B_t denotes an n-dimensional Brownian motion capturing

[6] The model setup also directly extends to the case of multiple experts. For details, see discussion in the end of the model setup.

normal-times idiosyncratic risk; and J_t is a (univariate) Poisson process with intensity λ. I refer to the Poisson arrival $dJ_t = 1$ as a tail event. In a tail event, the risky technologies suffer losses given by $\kappa + \sqrt{\phi} Z_t$. The vector of tail exposures κ captures the systematic exposure of each technology to the tail event, and $\sqrt{\phi} Z_t$ captures idiosyncratic risk. Idiosyncratic risk is normally distributed as $Z_t \sim \mathcal{N}(0, \Sigma)$, where $\Sigma = \sigma\sigma'$ is the n by n return covariance matrix during normal times. The scalar $\phi > 0$ controls the increase in idiosyncratic risk during a tail event. The yield of an asset is simply its expected return in the absence of tail events $\mu_\theta + \lambda\kappa$. In the model, reaching for yield takes place when experts tilt their portfolios toward assets with high yields rather than high expected returns.

Let μ_θ^+ be the maximum Sharpe ratio attainable by expert type θ, i.e., $\mu_\theta^+ = \sqrt{\mu_\theta' \Sigma^{-1} \mu_\theta}$, and $\kappa^+ = \sqrt{\kappa' \Sigma^{-1} \kappa}$ the maximum tail exposure. In the baseline analysis, we assume $\mu_\theta' \Sigma^{-1} \kappa = 0$, i.e., that tail risk does not carry a risk premium.[7]

Expert types. Experts come in two types. The *skilled* type ($\theta = S$) has an investment opportunity set with a Sharpe ratio $\mu_S^+ > 0$, and the *opportunistic* type ($\theta = O$) has a Sharpe ratio equal to zero ($\mu_O^+ = 0$). Investors observe the fund returns, but do not observe the expert type θ or the fund portfolio.[8] Their information can be represented by $\mathcal{F}_t^I = \{dR_s : s \leq t\}$. Because investors observe the fund return history continuously, tail events dJ_t are perfectly observable due to the discontinuous movement in fund returns they generate. Experts know their type and their information, which can be represented by $\mathcal{F}_t^E = \{(dR_s, \theta) : s \leq t\}$. At date 0, investors believe there is a probability P_t that the expert is *skilled* ($\theta = S$). I refer to P_t as the expert reputation.

Opportunities are scarce. The expert faces decreasing returns in their ability to invest in profitable opportunities. Let Q_t be a $1 \times n$ row vector of allocations that describe how the expert allocates capital across the n technologies, where a_t is the fund assets under management, and $\sigma_{r,t} = \sqrt{Q_t \Sigma Q_t^\top}$ is the fund return volatility. The expert faces transaction costs that increase with the total size of the fund portfolio, as assumed in Berk and Green (2004). In my model, portfolio size is captured by the fund asset volatility $a_t \sigma_{r,t}$. Specifically, I assume a cost function $c(a_t \sigma_{r,t}) > 0$, with $c'(\cdot) \geq 0$. More specifically, I assume in most of this paper that this cost function is quadratic. In Appendix B, I provide a micro-foundation for this cost function based on the assumption that each expert has access to a unique local market where they trade assets against mean-variance investors with hedging needs.[9]

Financial contracts. The expert raises capital from investors through a fund management contract. Here, I assume instantaneous contracts, where the expert quotes an intermediation fee $f_t \geq 0$ per dollar managed period by period. Further, as is standard in the principal-agent literature, I assume that when the expert is indifferent across portfolios, the expert will do the best thing for their investors. Specifically, I assume the expert places weight $m \geq 0$ on maximizing

[7] More precisely, μ_θ^+ is the maximum ratio of expected returns to the square root of the instantaneous quadratic variation (normal-times return volatility) across these technologies. We relax the assumption that tail risk has zero premium in Section 3.5.3.

[8] Investors can observe portfolios, as they do in mutual funds, as long as they do not distinguish individual assets' tail exposures and expected returns.

[9] This local market micro-foundation also motivates the heterogeneity in expert skill and the idiosyncratic risks that each expert faces.

the fund Sharpe ratio. Because of quadratic transaction costs, investors' utility is increasing in the fund Sharpe ratio.[10]

Equilibrium. I use perfect Bayesian equilibria as the equilibrium concept. Here, this concept has three implications: (i) investors take as given their own beliefs and the behavior of experts when choosing how much to invest; (ii) experts take as given the behavior of investors and investors' beliefs when choosing how to allocate capital; (iii) and investors' beliefs about the expert choices are consistent with expert choices, and investors update their beliefs about the type of expert according to Bayes' rule. I look for a stationary Markov equilibrium in which the only state variable is the expert's reputation P_t.

Discussion of the model setup. There are two key information frictions: I assume that the skilled expert has no credible channel to directly communicate either the quality of her investment opportunities or the tail exposures of her portfolio. The first assumption implies that investors must rely on an expert's past returns to learn about the expert type.[11] The assumption that the portfolio tail exposure cannot be observed implies that investors cannot adjust how they learn about an expert according to the tail exposure choice of the expert. Investors must instead rely on their own beliefs about the expert's portfolio choice.[12]

The existence of multiple technologies allows experts to change their tail exposure independently of fund volatility. Thus, volatility does not reveal the expert tail exposure. We will see that the opportunity set implicit in Equation (1) is spanned by two portfolios: one with the maximum Sharpe ratio and the other with the maximum tail exposure (per unit of volatility). However, the n-asset implementation allows a more transparent analysis of how the speed of capital flows depends on the asset tail exposure.

I focus on the case of a single expert, but the setup extends to the case where investors can observe and invest in multiple experts. Intuitively, the return history of other experts reveals no information about the shock history of a particular expert, because dB and Y shocks are idiosyncratic and the common tail event shock is observable by investors. Experts should be thought as investing in different set of assets or trading strategies (the shocks dB and Y are idiosyncratic) and this different opportunity set is the driver of their heterogeneity in expected returns μ_θ. The assets are different, but exposed to a common shock dJ_t which is a systematic shock hitting all asset markets simultaneously.

The key economic force driving an expert's portfolio decision is reputation concerns due to learning. The weight m is only useful for pinning down the expert's optimal portfolio when reputation concerns converge to zero. Theoretically, this only happens when $P_t = 1$, but because I solve the model numerically, a positive weight m is useful more broadly.[13] See Appendix D for additional discussion of the model setup.

---

[10] More broadly, one can think of m as the sensitivity to performance of compensation incentives. Recent evidence documented in Ma et al. (2019) shows that managers are often compensated on a risk-adjusted basis by fund management firms.

[11] While I rule out all forms of communication in the model, one can think of the initial expert reputation P_0 as the outcome of a "pre-trading" communication round. What is important for the model is that there is some residual uncertainty about the expert type after this communication round, so investors use returns to learn about the expert. The extensive empirical literature on mutual funds and hedge funds flows is consistent with this assumption.

[12] The opaqueness of the fund portfolio is motivated by the idea that the experts have a unique understanding of the assets they trade, and even though investors fully understand the environment, they are unable to evaluate expected returns or tail risks of specific portfolio positions.

[13] In section 2.5, I consider a case where I can solve the model analytically with $m = 0$.

Discussion of the model solution. Investors' learning behavior and the expert's portfolio policy are jointly determined in equilibrium. I choose to first solve the investor learning problem, and then later characterize the expert portfolio policy. This order is useful because it more transparently illustrates some general features of how the learning dynamic depends on the expert portfolio policy.

2.1. Learning

Let us first turn out attention to how investors learn about experts. Specifically, how should investors update their beliefs about the manager type θ, given an observed history of returns of an expert $\{dR_s, s \leq t\}$? Here is what they do: investors make conjectures about the equilibrium portfolio of each expert type, and then use differences in the implied conditional performance distributions to learn from the expert realized performance. In this section, I characterize investors' optimal learning policy, while I take as given investors' conjectures about equilibrium portfolios (Proposition 1). I then study the consequences for the learning dynamics of changes in these equilibrium portfolios. In particular, I show that learning is slower if the opportunistic expert reaches for yield, while learning is faster if it is the skilled expert that takes this action (Corollary 1.1).

Some general features of the equilibrium need to first be discussed. In the current setting, any equilibrium features a pooling of experts on observables. Opportunists can always mimic the skilled experts' choices, and never choose to separate from skilled types because that would imply immediate liquidation. This implies that observable quantities (portfolio volatility and fund assets) must depend only on reputation, and cannot depend on the expert type. Importantly, portfolio opaqueness implies that each expert can change the tail risk of their portfolio without affecting how investors interpret the expert performance.

Without loss of generality, we can decompose the expert portfolio choice in terms of a volatility choice and a portfolio composition choice for a given volatility. Specifically, portfolio Q_t can be decomposed as $Q_t = X_t \sigma_{r,t}$, where $\sigma_{r,t} = Q_t \Sigma Q_t^\top$ and $X_t \in \Omega = \{X \in \mathbb{R}^M | X \Sigma X^\top = 1\}$. Therefore the expert choice can be represented as a volatility choice $\sigma_{r,t}$ and a portfolio vector choice X_t with unit-variance. This decomposition is useful because it separates the component of the expert decisions that is observable by investors from the unobservable component. This decomposition implies that fund excess returns (gross of fees) per unit of volatility can be written as

$$dr_t^\theta \equiv \frac{Q_t dR_t^\theta}{\sigma_{r,t}} = X_t \left[(\mu_\theta + \kappa\lambda)\, dt - \kappa dJ_t \right] + dB_t + \phi^{\frac{1}{2}} z_t dJ_t - \frac{c(a_t \sigma_{r,t})}{a_t \sigma_{r,t}} dt. \tag{2}$$

Holding the expected return constant, a higher tail exposure increases the fund yield, i.e., the fund performance in any period without a tail event. For this reason, I refer to the fund yield $X_t(\mu_\theta + \lambda\kappa)$ as the fund normal-times performance. Because X_t is a unit-variance portfolio, it follows that $X_t \sigma dB_t = dB_t$ and $z_t = X_t Z_t$ are both univariate (standard) normals.

It is mathematically convenient to represent investors' beliefs in log-likelihood ratio space. Given a liquidation threshold $\underline{P}$, the log-likelihood distance to liquidation is $p = \log\left(\frac{P}{1-P}\right) - \log\left(\frac{\underline{P}}{1-\underline{P}}\right)$. I refer to both P and p as reputation.[14]

[14] The mapping from probability to log-likelihood is injective, and therefore without loss. I abuse notation and refer to functions that were defined in the probability space as if they were defined in the log-likelihood space.

The I superscript is used to denote investors' beliefs about a random variable. For example, I write $X^I(p,\theta) = E^I[X(p,\theta)]$ to denote investors' beliefs about the portfolio of an expert of type θ and reputation p.

A simple application of Bayes' rule in continuous time implies the following dynamics for investors' beliefs:

Proposition 1. ***Investor learning about expert skill.*** *Given investor beliefs $X^I(p,\theta)$, the expert reputation p_t satisfies,*

$$dp_t = g(p_t)\left(dr^\theta_{t-} - e(p_t)\,dt\right) + g_J(p_t)\phi^{-1}\left(dr^\theta_t - e_J(p_t)\right)dJ_t \tag{3}$$

for $p_t > 0$ and $dp_t = 0$ if there exists $s \le t$ such that $p_s \le 0$, with learning coefficients given by

$$g_J(p_t) = \left(X^I(p_t, O) - X^I(p_t, S)\right)\kappa, \tag{4}$$

$$g(p_t) = X^I(p_t, S)\mu_S - \lambda g_J(p_t), \tag{5}$$

$$e_J(p_t) = -\frac{1}{2}(X^I(p_t, S) + X^I(p_t, O))\kappa, \tag{6}$$

$$e(p_t) = \frac{1}{2}X^I(p_t, S)\mu_S - \lambda e_J(p_t). \tag{7}$$

Optimal learning in this environment consists of comparing the likelihood that a given return realization was generated by each expert type: $\theta \in \{S, O\}$. Learning takes place in both "normal times" and "tail events." In both states, the statistical learning problem investors face is how to differentiate between two normal random variables of identical variance but different means. The coefficients $g(p_t)$ and $g_J(p_t)$ track differences in expected performance across experts in each state $dJ \in \{0, 1\}$. For example, during normal times $g(p_t) \equiv \mathbb{E}^I_t\left[dr^S_{t-} - dr^O_{t-}\right]$, which is the difference in normal-times performance across types. Naturally investors learn more from returns when this signal-to-noise ratio is high. Return surprises are measured relative to how a p-reputation expert is expected to perform in each state. Specifically, $e(p)$ and $e_J(p)$ measure the expected average performance of a p-reputation expert. Because I am working in log-likelihood space, these expectations are a simple average across the types' expected returns in each state.

We obtain the speed investors learn in equilibrium by simply imposing that investors' beliefs must be consistent with experts' equilibrium portfolio policies ($X^I(p,\theta) = X(p,\theta)$),

$$\mathbb{E}_t[dp^S_t - dp^O_t] = (X(p,S)\mu_S)^2\left[\left(1 - \frac{g_J(p)\lambda}{X(p,S)\mu_S}\right)^2 + \frac{1}{\lambda\phi}\left(\frac{g_J(p)\lambda}{X(p,S)\mu_S}\right)^2\right], \tag{8}$$

where the ratio $g_J(p)\lambda/X(p,S)\mu_S$ captures the amount of "performance jamming" done by the opportunistic expert in equilibrium. This is the fraction of the performance advantage of the skilled expert that the opportunist makes up by taking more tail risk than the skilled type. When both types have the same tail exposure, $g_J(p) = 0$, their reputations diverge at the rate of the squared Sharpe ratio difference across types.[15] This learning speed is a natural benchmark, since it is the rate at which investors would learn if the opportunistic type were not strategic about his portfolio choice. Corollary 1.1 below starts by defining two basis portfolios, referred to as Sharpe and Tail portfolios, that together span the experts' investment opportunity set, and then

[15] It is simply the squared Sharpe ratio of the skilled type, since the opportunist Sharpe ratio is normalized to zero.

shows how the learning speed changes as each expert tilts their weight from the Sharpe to the Tail portfolio.

Corollary 1.1. ***Reaching for yield and the speed of learning.*** *Let the Sharpe portfolio be given by weights* $X_\mu = \frac{\Sigma^{-1}\mu_S}{\mu_S^+}$ *and the Tail portfolio be* $X_\kappa = \frac{\Sigma^{-1}\kappa}{\kappa^+}$*. Then it follows:*

(1) If both experts hold the Sharpe portfolio, the speed of learning is $(\mu_S^+)^2$*.*

(2) If both experts hold the Sharpe portfolio, a small increase in the opportunistic expert position in the tail portfolio decreases the speed of learning.

(3) If $\mu_S^+ > (\lambda + \frac{1}{\phi})\kappa^+$*, the skilled expert holds the Sharpe portfolio, and the opportunist expert holds the Tail portfolio, then the speed of learning increases in the skilled expert position on the Tail portfolio.*

Corollary 1.1 above starts by defining the two basis portfolios. I will show that the experts' optimal portfolio choice is always a combination of these two portfolios. While I have not characterized the optimal behavior yet, Result (2) in Corollary 1.1 above shows that the opportunist can reduce the speed of learning by increasing performance in states when informativeness is high (normal times) at the expense of lower performance when informativeness is low (tail events), i.e., reach for yield by the opportunist reduces the speed of learning. Result (3) in Corollary 1.1 focuses instead on the skilled-type tail exposure choice. This result implies a trade-off between the (static) allocation of capital and the speed of learning. There is a trade-off here because the skilled expert can increase the speed of learning by reaching for yield, which reduces expected returns and makes capital allocation worse in the short term. However, we will see that faster learning enables capital to flow faster to the skilled expert, who allocates capital better than the opportunistic type. This trade-off between the static and dynamic allocation of capital is one of the key trade-offs in this paper.

2.2. Financial experts

I now solve the expert portfolio allocation problem. The key goal of this section is to characterize experts' optimal portfolio in terms of a pair of incentive functions (Propositions 3 and 4). Here, I take as given the learning dynamics and the capital supply policies that investors follow.[16] The financial expert maximizes the net present value of their compensation flow plus the weight m they place on maximizing their portfolio Sharpe ratio. The expert chooses the fee policy f and the portfolio policies σ_r and X:

$$\sup_{f,\sigma_r,X} \mathbb{E}_u^E \left[\int_u^\infty e^{-\rho(t-u)} \left(m X_t \mu_\theta + a_t f_t \right) dt \right] \quad s.t. \quad X_t \in \Omega,\ f_t \geq 0, \sigma_{r,t} > 0. \tag{9}$$

The nonnegative fee constraint implies the expert cannot make payments to investors to delay/avoid liquidation. The experts' problem depends on investors' beliefs p_t because they shape investors' decision of how much capital to allocate to the expert. It is the investors' capital allocation decisions that makes the experts' problem dynamic. Thus, the Hamilton–Jacobi–Bellman equation for the expert can be written as,

[16] Investors' capital supply policies are characterized in Section 2.4.

$$
\begin{aligned}
&(\rho+\lambda)V(p_t,\theta)\\
&=\sup_{\{X,\sigma_r,f\}\in\Omega\times\mathbb{R}_+^2} mX\mu_\theta + a_t f + V_p(p_t,\theta)g(p_t)\left(X(\mu_\theta+\lambda\kappa)-e(p_t)\right) \qquad (10)\\
&\quad+\frac{1}{2}V_{pp}(p_t,\theta)g^2(p_t)+\lambda\mathbb{E}_z\left[V\left(p_t-\frac{g_J(p_t)}{\phi}\left(X\kappa+\sqrt{\phi}z+e_J(p_t)\right),\theta\right)\right],
\end{aligned}
$$

for $p_t \geq 0$, with the boundary condition $V(p,\theta)=0$ for any $p \leq 0$. The expectation $\mathbb{E}_z[]$ is taken over the standard normal random variable z.

On the left—hand side, we have the effects of discounting due to the passage of time (ρ) and the arrivals of tail events (λ). The first two terms on the right—hand side (RHS) of the first line capture the instantaneous compensation flow. The third term captures the valuation effects of reputation growth during normal times. On the second line, we first have the effects of reputation risk during normal times. The last term captures what happens with reputation during a tail event: the expert's reputation jumps to a new and uncertain value because of the lumpy amount of information revealed during a tail event.

The portfolio choice X appears in the compensation term, in the normal-times reputation growth term, and on the tail-event term. Importantly, it does not show up on the investors' learning coefficients, which only depend on the expert's equilibrium choice.

A benchmark: portfolio choice when the fund is transparent. In this case, investors can condition on the experts' tail exposure and the opportunistic expert pools with the skilled type in order to remain undetected. Both types have the same tail exposure, thus tail event performance is uninformative, and the difference in normal-times performance across types solely reflects each expert's expected returns. Formally, transparency implies $g_J(p_t)=0$ and $e_J(p_t)=-X\kappa$. Rearranging the first-order condition with respect to the portfolio choice, substituting both the learning coefficients and the definition of the Sharpe portfolio, I obtain

$$X^* = X_\mu \times \mu_S^+(m+V_p(p,S)g(p))/\zeta, \qquad (11)$$

where ζ is the Lagrange multiplier associated with the unit-variance constraint.[17] Proposition 2 summarizes the equilibrium in this case[18]:

Proposition 2. ***Learning speed when portfolios are transparent.*** *If $X_t^\top\kappa \in \mathcal{F}_t^I$, then (i) $X(p,S)=X(p,O)=X_\mu$, (ii) $\mathbb{E}[dp_t^S-dp_t^O]=(\mu_S^+)^2$.*

The intuition for this result is as follows. Equation (11) above implies the optimal portfolio is proportional to the Sharpe portfolio. Because reputation concerns $V_p(p_t,S)g(p_t)$ are weakly positive in equilibrium, the expert has always a long position in the Sharpe portfolio. This implies that her optimal portfolio is simply the Sharpe portfolio given that her portfolio choice X must also have unit variance.[19]

Portfolio choice when the fund is opaque. When investors do not observe the fund tail exposure, the learning coefficients only depend on investors' beliefs about experts' portfolios, and the opportunistic expert no longer has to pool with the skilled type.

[17] Note that the expert is free to choose a portfolio of any volatility, but because volatility is observable by investors while the rest of the portfolio is not, it is important to deal with these two decisions separately. The choice of portfolio volatility is studied in Section 2.4.

[18] See Corollary 2.1 in Appendix A for an analytical characterization of the experts' value functions in this case.

[19] Note that we study the choice of portfolio volatility in Section 2.4.

Rearranging the first-order condition associated with Equation (10), and omitting the dependence on p_t and θ for expositional convenience, I obtain that the optimal portfolio is a weighted average of the Sharpe and Tail portfolios,

$$X^* = X_\mu \mu_\theta^+ \frac{(m + V_p g)}{\zeta} + X_\kappa \lambda \kappa^+ \frac{V_p g - \frac{g_J}{\phi}\mathbb{E}_z\left[V_p\left(p_{t+}, \theta\right)\right]}{\zeta}, \tag{12}$$

where p_+ denotes the reputation of the expert after a tail event. The weight on the Sharpe portfolio consists of the performance weight m and the normal-times reputation concerns $V_p g$. The weight on the Tail portfolio is simply the net effect of reputation concerns during normal times and tail events. Normal-times concerns push the expert to reach for yield and increase the weight on the tail portfolio, but reputation concerns during a tail event push this weight down. The expert will reach for yield aggressively whenever normal-times concerns dominate. This happens when the expert's reputation is low. Intuitively, the expert will be more concerned with her short-term reputation evolution when she is close to liquidation.[20]

Focusing on the extreme cases is useful for intuition. If tail-event volatility $\phi \to \infty$ and the weight $m \to 0$, we have that the expert chooses $X^* \propto \mu_\theta^+ X_\mu + \lambda\kappa^+ X_\kappa$, which is the portfolio that maximizes the fund yield. Conversely, if the expert has identical reputation concerns across normal times and tail events, then $V_p g = \frac{g_J}{\phi}\mathbb{E}^E\left[V_p\left(p_{t+}, \theta\right)\right]$, and the skilled expert portfolio policy is to maximize expected returns, $X^* = X_\mu$. It is the wedge in reputation incentives between normal times and tail events that drives the expert to reach for yield.

Equation (12) characterizes the optimal choice implicitly as the solution of an n-dimensional fixed point problem. Intuitively, the expert's portfolio influences her expected reputation after a tail event $p_{t+} = p_t - \frac{g_J}{\phi}\left(X\kappa + \sqrt{\phi} z + e_J\right)$, and the change in reputation impacts her reputation concerns because of the concavity of her value function. This feeds back into her portfolio choice. In the reminder of this section, the problem of each expert separately analyzed, in order to further characterize optimal behavior.

Skilled portfolio. I will start by showing that the skilled expert's multi-dimensional portfolio choice can be represented with a scalar y that summarizes the expert's net investment incentives.

Lemma 1. ***Scalar representation of the skilled portfolio.*** *For any optimal portfolio X for the skilled expert, there is a scalar y such that $\mathcal{X}(y) = X$, where*

$$\mathcal{X}(y) \equiv \frac{\mu_S^+}{\sqrt{(\mu_S^+)^2 + (\lambda\kappa^+)^2 y^2}} X_\mu + \frac{\lambda\kappa^+ y}{\sqrt{(\mu_S^+)^2 + (\lambda\kappa^+)^2 y^2}} X_\kappa. \tag{13}$$

The denominator in Equation (13) rescales weights so that $\mathcal{X}(y)$ is unit-variance. The scalar y can be interpreted as the expert reach-for-yield incentive, because it measures how much the expert values an increase in yields relative to expected returns. If the expert only cares about expected returns, then $y = 0$. If she only cares about yield, then $y = 1$. If she only cares about tail performance, then $y = \infty$. Going forward I represent the skilled choice and investors' conjectures about this choice in terms of y and y^I.[21]

---

[20] I will show that in equilibrium $V()$ is increasing and concave, $g(p_t) \geq 0$, and $g_J(p_t) \geq 0$; thus an increase in the portfolio tail exposure increases reputation growth during normal times at the cost of a reduction in reputation growth during tail events.

[21] There is a one-to-one mapping between incentives y and the skilled expert portfolio.

I now substitute out the learning coefficients in Equation (12) and represent the optimal weights on the two basis portfolios as a function of incentives y, investors' beliefs about these incentives y^I, and investors' beliefs about the difference in tail exposures across types $g^I_J(p) = X^I(p, O)\kappa - \mathcal{X}(y^I)\kappa$. I then construct the function that characterizes the expert incentives y by collecting the terms that multiply the Sharpe portfolio X_μ and the Tail portfolio X_κ, and normalizing the latter by the former.

Proposition 3. ***Skilled expert portfolio choice.*** *For a given reputation p and investors' beliefs about the experts' portfolios, g^I_J and y^I, the skilled expert incentives solves $y = \mathcal{D}_S(p, y|g^I_J, y^I)$, where $\mathcal{D}_S$ is*

$$\mathcal{D}_S(p, y|g^I_J, y^I) = \frac{(\mathcal{X}(y^I)\mu_S - \lambda g^I_J)V_p - \lambda \frac{g^I_J}{\phi}\mathbb{E}_z\left[V_p\left(p - \frac{g^I_J}{\phi} \times \left(\mathcal{X}(y)\kappa - \frac{2\mathcal{X}(y^I)\kappa + g^I_J}{2} + \sqrt{\phi}z\right), S\right)\right]}{m + (\mathcal{X}(y^I)\mu_S - \lambda g^I_J)V_p}. \tag{14}$$

The function $\mathcal{D}_S(\cdot)$ summarizes the expert marginal investment incentives as a function of investors' beliefs and the expert's own portfolio choice. Concavity of the value function (a property we show in Section 2.5) implies that the marginal value of reputation in a tail event is increasing in reach-for-yield incentives y. Under certain conditions, this property implies that there is a solution to the fixed point problem in Proposition 3. However, in the interest of space, I will prove this formally once I impose consistency between experts policies and investors' conjectures about these policies in Section 2.3.

Proposition 3 is a useful place to discuss the role played by compensation incentives m. From Equation (14) we see that m shrinks reach-for-yield incentives toward zero. Qualitatively, this plays a role when the expert reputation is large and the expert value function sensitivity to reputation approaches zero. In this case, an $m = 0$ implies that $\mathcal{D}_S$ would be undefined in the limit $P = 1$, and very poorly behaved numerically for P close to 1. This is the reason I use $m > 0$ in the numerical analysis. A $m > 0$ induces the expert to maximize the fund Sharpe ratio when reputation concerns become weak.

Opportunist portfolio. In contrast with the skilled expert, the opportunist does not face a trade-off between expected returns and yield. Because his Sharpe ratio is zero, he can substitute performance between normal times and tail events without any impact on the fund's expected returns.[22] Solving the FOC in Equation (12) for the opportunist portfolio, I obtain that the opportunist portfolio is proportional to the Tail portfolio,

$$X^* = X_\kappa \mathbb{E}^E\left[g(p)V_p(p, O) - \frac{h}{\phi}V_p(p_+, O)\right]\lambda\kappa^+/\zeta, \tag{15}$$

where the proportionality term is again the expert (net) reputation concerns. When reputation concerns are positive, he reaches for yield as much as he can.[23] When the investment opportunity set is very rich, i.e., κ^+ large, this constraint does not bind, as the opportunist reaches for yield

[22] In Section 3.5.3 I consider the case that tail-risk premium is not zero. I find that a positive risk premium further increases the reach-for-yield incentives of the expert.

[23] In the next section, I demonstrate that reputation concerns are never negative in equilibrium.

just enough to balance out normal times and tail event reputation incentives. In this case, the zero reputation concerns condition ($\mathcal{D}_O(\cdot) = 0$) pins down his optimal portfolio. Proposition 4 below formalizes this discussion.

Proposition 4. ***Opportunist portfolio.*** *Given investors' beliefs y^I and g_J^I, the optimal portfolio policy of the opportunistic expert is given by*

$$g_J = \begin{cases} -\kappa^+ - \mathcal{X}(y^I)\kappa & \hat{g}_J + \mathcal{X}(y^I)\kappa < -\kappa^+ \\ \hat{g}_J & \hat{g}_J + \mathcal{X}(y^I)\kappa \in [-\kappa^+, \kappa^+] \\ \kappa^+ - \mathcal{X}(y^I)\kappa & \hat{g}_J + \mathcal{X}(y^I)\kappa > \kappa^+, \end{cases} \tag{16}$$

where $\hat{g}_J$ solves $\mathcal{D}_O(p, \hat{g}_J | g_J^I, y^I) = 0$. The incentive function $\mathcal{D}_O(\cdot)$ is given by

$$\begin{aligned} \mathcal{D}_O(p, g_J | g_J^I, y^I) = &\left(\mathcal{X}(y^I)\mu_S - \lambda g_J^I\right) V_p \\ &- \lambda \frac{g_J^I}{\phi} \mathbb{E}_z \left[V_p \left(p - \frac{g_J^I}{\phi} \left(g_J - \frac{g_J^I}{2} + \sqrt{\phi} z \right), O \right) \right]. \end{aligned} \tag{17}$$

2.3. Equilibrium

In sections 2.1 and 2.2 we have studied separately the investor learning problem and the expert portfolio problem. I now characterize equilibrium by requiring that investors' beliefs about experts' portfolio policies coincide with the optimal policies of each expert.

After characterizing equilibrium in Proposition 5 below, I derive three properties of the equilibrium. First, the skilled expert always reaches for yield (Corollary 5.1); second, reputation is increasing in performance (Corollary 5.2); and third, learning is slower when the fund is opaque (Corollary 5.3).

Proposition 5. ***Equilibrium portfolios.*** *For any $p \geq 0$, the portfolio policies $\{g_J(p), y(p)\} : \mathcal{R}^+ \to [0, \kappa^+] \times [0, 1]$ are consistent with equilibrium if they solve the following system of equations:*

$$\mathcal{D}_S(p, y | g_J, y) = y, \tag{18}$$

$$\mathcal{D}_O(p, g_J | g_J, y) \left(\kappa^+ - (g_J + \mathcal{X}(y)\kappa)\right) = 0, \tag{19}$$

$$\mathcal{D}_O(p, g_J | g_J, y) \geq 0, \tag{20}$$

$$g_J + \mathcal{X}(y)\kappa \leq \kappa^+. \tag{21}$$

The equilibrium restriction imposes the requirement that the incentives have to be self-generating: investors' beliefs about experts' portfolio policies must in turn create incentives for these policies to be optimal for each expert type. Equation (18) imposes this restriction on the skilled expert and Equations (19)-(21) on the opportunistic type.[24]

Fig. 1 illustrates the two types of equilibrium consistent with Proposition 5. In Panel A, we have the equilibrium where Equation (21) is slack and Equation (20) holds with equality. In this case, the opportunistic expert is in an interior optimum where his portfolio choice just balances

[24] I need multiple conditions to characterize the opportunist portfolio, because his portfolio choice is sometimes at the corner of his investment opportunity set.

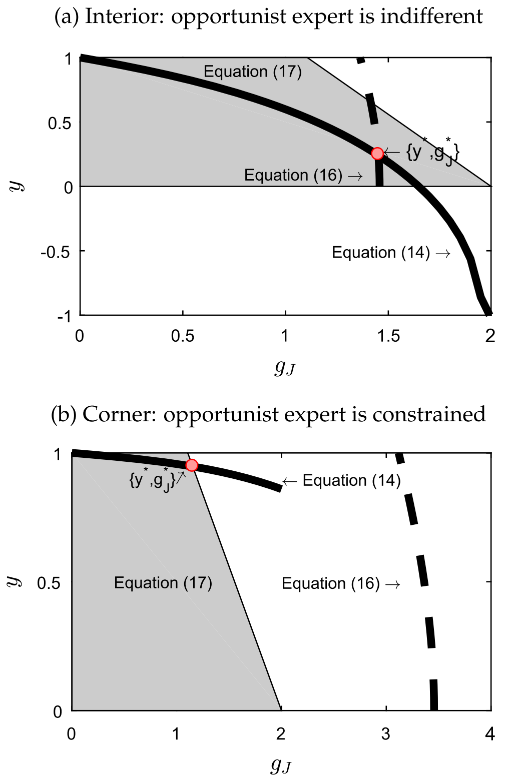

Fig. 1. **Equilibrium**. Panels (a) and (b) show equilibrium determination according to Proposition 5. Panel (a) shows the case where Equation (21) is slack (the grey area) and the equilibrium is determined by the opportunistic expert indifference condition [Equation (20)]. Panel (b) illustrates the case where Equation (21) binds and Equation (21) is slack (the area at the left of the dashed line). In this case the opportunistic expert would like to take on more tail risk than his investment opportunity set permits.

out the competing reputation incentives. In Panel B, we have the equilibrium where Equation (21) holds with equality and Equation (20) is slack. In this case the opportunistic type is in a corner and reaches for yield aggressively. His investment opportunity set does not allow him to take more tail risk.

To gain more insight on the forces determining the equilibrium portfolio, I discuss the skilled expert equilibrium portfolio choice y, taking as given the opportunist equilibrium response g_J; I then discuss the opportunistic expert equilibrium response g_J, while taking y as given.

Skilled experts. Let us look at the fixed-point Equation (18) for the case in which the weight m is zero:

$$y = 1 - \left(\frac{\mathcal{X}(y)\mu_S - \lambda g_J}{\lambda g_J/\phi} \frac{V_p(p,S)}{\mathbb{E}\left[V_p\left(p_+,S\right)\right]} \right)^{-1}. \tag{22}$$

The ratio $\frac{V_p(p,S)}{\mathbb{E}[V_p(p_+,S)]}$ is referred to as the expert short-term bias. If the short-term bias is above 1, the expert places a higher value on an increase in reputation in normal times than in a tail event. The higher this ratio, the more willing the expert is to sacrifice tail-event performance to improve normal-times performance, i.e., reach-for-yield incentives y are stronger. Multiplying this ratio we have $\frac{\mathcal{X}(y)\mu_S-\lambda g_J}{\lambda g_J/\phi}$, which I refer to as the informativeness ratio because it is the signal-to-noise ratio of normal-times performance relative to the signal-to-noise of tail-event performance. A high informativeness ratio also increases y, as reputation incentives depend not only on how much the expert cares about her reputation, but also on how much performance impacts reputation in each state. Note that when $g_J = 0$, reach-for-yield incentives are at a maximum $y = 1$, and that initially y decreases with g_J, as tail-event performance becomes progressively more informative than normal-times performance. This reduces the incentives to reach for yield. However, note that an increase in g_J implies that in a tail event the reputation of the skilled expert is expected to increase by more, leading eventually to a reduction in concerns with reputation risk during a tail event.

Corollary 5.1 shows that, in equilibrium, the incentives y is strictly above zero, i.e., the skilled type always reaches for yield in equilibrium. The result below also shows that for a given opportunist policy g_J, the skilled policy is unique.

Corollary 5.1. ***The skilled expert always reaches for yield.*** *Let $g_J = cte \geq 0$, $V_p(p, S) > 0$, $V_{pp}(p, S) < 0$, and ϕ high enough. Furthermore, let $\mu_S^+ > \lambda\kappa^+(1 + \frac{1}{\phi})$ then for each p, then it exists a unique $y(p|g_J) \in (0, 1]$ that satisfies Equation (18).*

Opportunistic expert. Equation (19) imposes one of two restrictions on the opportunist policy g_J: either he must be in a corner and choose tail exposure equal to κ^+, or his policy g_J is such that reputation incentives balance out,

$$(\mathcal{X}(y)\mu_S - \lambda g_J)\, V_p = g_J\lambda/\phi\mathbb{E}\left[V_p\left(p_+, O\right)\right]. \tag{23}$$

If the expert cares equally about reputation in both states, he chooses a tail exposure that equates the informativeness of performance across them, which is also the choice that minimizes the rate of learning. However, if he has a strong short-term bias, he will choose to lose more reputation in a tail event so he loses less in normal times. This interior solution is not feasible when the expert's short-term bias is sufficiently strong, as this balance of reputation incentives might require a tail exposure higher than κ^+. This behavior also implies that reputation is always increasing in performance.

Corollary 5.2. ***Reputation is increasing in performance.*** *Let $V(p, O)$ be increasing and concave; we have that $g(p) = \mathcal{X}(y) - \lambda g_J(p) \geq 0$ and $g_J(p) \geq 0$.*

The proof is intuitive. If g_J were negative, it would imply that reputation is increasing in performance during normal times and decreasing in performance during tail events. This would allow the opportunist to increase reputation growth state-by-state by increasing his portfolio tail exposure. Conversely, g_J cannot be so large that reputation decreases in normal-times performance, because now the opportunist would increase reputation growth by reducing the portfolio tail exposure. Therefore such a large g_J cannot be an equilibrium, either. The proof highlights the importance of the opportunistic expert strategic behavior in ruling out equilibria where the relationship between performance and reputation is inverted.

Together with Corollary 5.1 and Equation (8), this result implies that learning is always slower when the fund portfolio is opaque versus when the fund is transparent. Intuitively, the ability to take on tail risk in a hidden fashion allows the opportunistic type to manipulate performance to delay learning.

Corollary 5.3. ***Learning is slower when portfolios are opaque.*** *If $\mu_S^+ > \lambda\kappa^+$ and $\phi > 1/\lambda$, then $\mathbb{E}[dp_t^S - dp_t^O] \leq (\mu_S^+)^2$.*

In Lemma 2 below, I show that if tail-event returns are sufficiently volatile, there always exists a corner equilibrium, as depicted in Fig. 1(b).

Lemma 2. ***Equilibrium existence.*** *Let the conditions of Corollary 5.1 hold; then for any $p \geq 0$, there exists at least one pair $\{g_J, y\}$ consistent with equilibrium.*

Note that learning will be even slower in an equilibrium where g_J is interior, because in this case the opportunist is minimizing the speed of learning.[25]

Strategic complementaries and multiplicity. The fixed point problem in Proposition 5 features strategic complementaries between investors' beliefs and experts' portfolio decisions. To show this clearly, I consider the extreme case where there is no learning during tail events ($\phi \to \infty$) and the weight m is zero. The opportunist wants to match the normal-times performance of the skilled type because there is no cost associated with performing poorly during tail events, i.e., he wants to choose $g_j = \lambda^{-1} X(p, S)\mu_S$. Let's first conjecture that the skilled expert does not reach for yield $y = 0$. If $\kappa^+ \geq \lambda^{-1}\mu_S^+$, then the opportunist target choice is feasible for $y = 0$. Performance differences across types during normal times disappear. Performance ceases to have information, the skilled type has no reputation incentives, and our initial conjecture that she does not reach for yield ($y = 0$) is correct in equilibrium. Now let's conjecture that she reaches for yield aggressively ($y = 1$); then because $\mathcal{X}(y = 1)(\mu_S + \lambda\kappa) > \lambda\kappa^+$, the opportunistic type can no longer match her performance. This results in normal-times performance being informative and consequently $y = 1$ is an optimal choice for the skilled expert. So both $\{y = 0, g_J = \mu_S^+/\lambda\}$ and $\{y = 1, g_J = \kappa^+ - \mathcal{X}(y = 1)\}$ are an equilibrium of this economy. This example illustrates how equilibrium multiplicity emerges from the two-way interaction between investors and experts.

Equilibrium selection. I have shown that the model delivers a trade-off between the speed of learning and the efficiency that capital is allocated. The existence of multiple equilibria implies that some equilibrium will feature faster learning and have a less efficient allocation of capital, i.e., a higher y, and other equilibria will feature slower learning but a more efficient allocation of capital, i.e., a lower y. The purpose of this paper is to highlight the link between reach-for-yield behavior and slow-moving capital, so I choose to focus on the equilibrium that features *faster* learning and *faster* capital flows (higher y). The formal equilibrium selection criteria are stated in Appendix A.

2.4. Capital flows

I now build on the equilibrium portfolio choice and learning dynamics we have studied so far to characterize the dynamics of capital flows.

Investors allocate capital to a given expert's as long as the expert reputation is high enough—i.e., their investment policy is of the threshold type. How much they invest depends on how quickly the expert profitability decreases with scale. This effect is captured by the cost function

[25] Formally, in the interior case Equation (23) shows that the opportunistic expert minimizes the speed of learning adjusted for reputation concerns. Only when the expert places the same value on increases in reputation during normal times and tail events, he minimizes the speed of learning.

$c(\cdot)$. First, note that the threshold-type policy generates fund liquidations in equilibrium. This feature of the environment is quite general and emerges naturally if investing with the opportunistic type is costly.[26]

For the intensive margin of the capital allocation decision, all that is required to connect the speed of learning and the speed of capital flows is that investors allocate more capital when they expect higher returns. We capture this property by assuming the expert faces quadratic trading costs that are increasing in the total volatility of the expert portfolio, i.e., $a_t\sigma_{r,t}$. We provide a micro-foundation for this cost function in Appendix B, based on the assumption that experts trade against mean-variance traders in these asset markets. An alternative micro-foundation is that fund investors themselves have mean-variance preferences with respect to their investments in the fund.[27] As in Berk and Green (2004), the capital allocated to an expert is pinned down by investors' zero profit condition, which states that the fund returns net of costs should be equal to the charged fees,

$$P_t\sigma_{r,t}X(P_t,\theta)\mu_S - \sigma_{r,t}c(a_t\sigma_{r,t}) = f_t. \tag{24}$$

The equilibrium determination of the capital allocated to an expert works through the cost function. As investors allocate more dollars to the expert, costs increase and profitability decreases. Investors allocate more dollars to the expert until the condition (24) holds and net-of-cost returns equate to the paid fees. The "decreasing returns to scale" hypothesis implied in the increasing shape of the cost function c is rooted in the idea that the expert skill is scarce. Intuitively, as more capital is deployed to chase an asset pricing anomaly, the returns erode as the anomaly disappears. Implicit in the assumption that returns are decreasing at the fund level is that the expert has an strategy that is truly unique to her. In the micro-foundation provided in Appendix B this uniqueness is related to the set of markets that she has access to. For a broader discussion of the empirical evidence for this "decreasing returns to scale" hypothesis see Pástor et al. (2015) and Berk and Van Binsbergen (2015). Equation (24) pins down the expert assets and the reputation threshold at which investors liquidate the expert. I summarize the liquidation policy in Proposition 6 below.

Proposition 6. ***Expert liquidation.*** *There is a reputation cutoff $\underline{P}$ such that the fund is liquidated permanently the first time $P_t < \underline{P}$. Let the fee charged by the expert at liquidation be zero, then*

$$\underline{P} = \frac{c(0)}{X(y=0,S)\mu_S}. \tag{25}$$

Assuming that the optimal fee at liquidation is zero (which we prove below), the costs $c(0)$ are exactly compensated by the expected return of the expert at the liquidation threshold. The expert is liquidated for any reputation below this threshold because fees cannot be negative and perceived expected returns are decreasing in reputation.

[26] In environments with perfect competition (as in this paper) this happens the first time the expected return of the expert (net of the minimum fee) becomes negative. In an environment with less competition among investors, they capture some of the upside of investing with the expert and would optimally delay liquidation, but the optimal policy would still involve eventual liquidation. Standard forms of risk-aversion would also lead to this threshold property because investors with experts close to liquidation would have only a small allocation with the expert (since perceived expected returns must be low close to liquidation), and would therefore behave as risk-neutral investors with respect to the liquidation decision.

[27] I choose the first interpretation because what is scarce here is the expert skill, and not risk-bearing capital, as the risk-aversion interpretation would imply.

We now turn to the intensive margin of the capital allocation decision. Given a quoted fee, investors decide how much capital they will allocate to the expert. Experts internalize this behavior in their fee decision. I start by substituting the investor investment policy [Equation (24)] into the Bellman equation [Equation (10)]. I then take the first-order condition with respect to the fund asset volatility $\sigma_{a,t} = a_t\sigma_{r,t}$, and obtain

$$P_t X(p_t, S)\mu_S = c'(\sigma_{a,t}). \tag{26}$$

At the optimum, the increase in dollar returns due to an extra dollar invested (LHS) must be equal to the increase in costs (RHS). Only the fund asset volatility is determined here, because there are no borrowing frictions. In a world with leverage constraints, experts would favor small fees to attract more capital and avoid taking leverage.

Proposition 7. ***Equilibrium supply of capital.*** *Let $\sigma_a(p_t, y_t) = a_t\sigma_{r,t}$ be the fund capital, $f_a(p_t, y_t) = a_t f_t$ the expert compensation, and $\mathbb{P}(p) = \frac{\underline{P}e^p}{1-\underline{P}+\underline{P}e^p}$ be the function that maps p into the probability that the expert is skilled. It follows that given reputation p_t, incentives y_t, and cost function $c(x) = \psi_0 x + \frac{\psi_1}{2}x^2$, experts' and investors' optimality implies $\sigma_a(p_t, y_t) = \frac{\mathbb{P}(p_t)\mathcal{X}(y_t)\mu_S - \psi_0}{\psi_1}$ and $f_a(p_t, \pi_t) = \frac{\psi_1}{2}\sigma_a(p_t, y_t)^2$.*

From slow learning to slow-moving capital. I have shown that there is a trade-off between the speed of learning and the optimal allocation of capital (Corollary 1.1). I now show how the reduction in the speed of learning slows down the flow of capital (Corollary 8.1), and characterize the conditions under which the skilled expert capital allocation decision faces a trade-off between static and dynamic efficiency in the allocation of capital (Corollary 8.2). Here I focus on the rate at which capital flows to the skilled expert. In Section 3.4 I consider asset level flows as well.

In the Proposition 8 below, I fix the skilled expert reach-for-yield incentives y, and focus on an approximation that allows for an analytical characterization of the link between reach-for-yield behavior and the speed of capital flows. In Section 3.3, I show numerically that this connection holds up once we take into account the response of reaching-for-yield incentives to the change in reputation.

Proposition 8. ***Capital flows with opaque and transparent funds.*** *The speed that capital flows when the fund is transparent:*

$$E\left[d(\sigma_a^S(p,y) - \sigma_a^O(p,y))|y = g_J = 0\right] = \mathbb{P}'(p)\frac{\mu_S^+}{\psi_1}(\mu_S^+)^2. \tag{27}$$

When the fund is opaque we have

$$\begin{aligned} &E\left[d\left(\sigma_a^S(p,y) - \sigma_a^O(p,y)\right)|y\right] \\ &\approx \mathbb{P}'(p)\frac{\mathcal{X}(y)\mu_S}{\psi_1}\left((\mathcal{X}(y)\mu_S - g_J(y)\lambda)^2 + \lambda\frac{g_J(y)^2}{\phi}\right). \end{aligned} \tag{28}$$

The expressions (27) and (28) are intuitive. The term outside the brackets is the sensitivity of fund capital to reputation. The term inside the brackets is the speed of learning. When $y = 1$, the sensitivity of fund capital to reputation is the same across the two information environments, but the speed of learning will be different as long as $g_J > 0$, which is always the case in equilibrium.

Corollary 8.1. ***Capital flows more slowly when the fund is opaque.*** *If ϕ is sufficiently high, fund opaqueness reduces the speed of capital flows.*

Corollary 8.1 shows that the reach-for-yield by the opportunistic type reduces the speed of capital flows when the fund is opaque. I show below that reach-for-yield behavior by the skilled expert increases the speed at which capital flows, but does so at the cost of lower expected returns. By distorting her portfolio toward high-yield assets, the skilled expert increases her normal-times performance and the speed of capital flows. These results are central to this paper. A trade-off between present and future capital allocation emerges as the skilled expert has to choose between the optimal allocation of capital in a given time, which is consistent with $y = 0$, and the future allocation, which improves faster when she reaches for yield $y > 0$.[28]

Corollary 8.2. ***Capital flows faster when the skilled expert reaches for yield.*** *If $\mu_S^+ > (\lambda + \frac{1}{\phi})\kappa^+$ and $g_J(y) = \kappa^+ - \mathcal{X}(y)\kappa$, i.e., the opportunist is in a corner, then the speed of capital flows increases in y at $y = 0$.*

2.5. Analytical solution

In this section I focus on a case where the experts' value functions can be solved analytically. This enables some properties of the solution to be proven formally.[29] The solution is presented in Proposition 9 and it is discussed below:

Proposition 9. ***Equilibrium valuations: an analytical solution.*** *Let $m = 0$, $\phi \to \infty$, and $a_t f_t = \overline{f}$; then*

$$V(p, \theta) = \frac{\overline{f}}{\rho}\left(1 - \exp\left(-\iota_\theta \times p\right)\right), \tag{29}$$

where $\iota_S = \frac{\sqrt{1+\frac{8\rho}{g^2}}+1}{2}$, $\iota_O = \frac{\sqrt{1+\frac{8\rho}{g^2}}-1}{2}$, *and* $g = \sqrt{(\mu_S^+)^2 + (\lambda\kappa^+)^2} - \lambda\kappa^+$.

The key for obtaining an analytical solution is that lack of learning during tail events ($\phi \to \infty$) implies $y = 1$ and $X_O\kappa = \kappa^+$, and transforms the integro-differential equation in (10) into an ordinary differential equation (ODE). Finally, the assumption of constant scale and constant fee makes this ODE have constant coefficients.

In Equation (29) we see that the expert's valuation increases at a decreasing rate, with the valuation reaching the value of the discounted fees when the probability of liquidation is zero ($p \to \infty$). In this case, valuation changes only because the probability of fund liquidation changes with reputation. The coefficient ι_θ summarizes how fast the liquidation risk decays with the experts' reputation. The key determinant of the slope coefficient is the speed of learning. Because returns are uninformative during tail events, the speed of learning is given by g, the difference in normal-times performance between experts. We see that $\iota_S > \iota_O$, so liquidation risk decays faster for the skilled type. Intuitively, because the opportunist's reputation drifts toward

[28] Capital allocation improves faster when $y > 0$, because the skilled expert allocates capital better than the opportunistic expert.

[29] I verify numerically in the Section 3 that these properties hold more generally.

zero, an increase in reputation only produces a temporary delay in liquidation. This dynamic contrasts with the skilled type, for whom an increase in reputation leads to a permanent reduction in the probability of liquidation because her reputation trends upward. Thus, the heterogeneity in reputation concerns is a result of the endogenous learning dynamics.

Proposition 9 above gives us three properties of the experts' value function that are key to the results in Section 2.3. Specifically, we learn that value functions are increasing in reputation and concave, and that the skilled expert's value function is more concave. In the next section, I show that these properties hold more generally.

3. Analysis

In this section, a numerical example is used to expand on the model's implications. I solve for $V(p, \theta)$ numerically using finite-difference methods (Appendix C provides details). The baseline parameter values are given below. I choose a parameter combination to illustrate the model's predictions.[30]

Numerical example: $\mu_S^+ = 1$, $\mu_S \propto 1_n$, $\kappa^+ = 2$, $\mu_S^\top \Sigma\kappa = 0$, $n = 9$, $\lambda = 0.25$, $\sqrt{\phi} = 3$, $\rho = 0.05$, $\psi_0 = 0.5$, $\psi_1 = 0.5$, $m = 0.2$.

3.1. Expert valuation dynamics and incentives to reach for yield

The equilibrium reach-for-yield incentives are the result of dynamic incentives reflected on experts' valuations. Figs. 2(a) and 2(b) show how valuations change with the experts' reputation. Valuations increase with reputation as the probability of liquidation decreases. Valuations are higher for the skilled type because her performance is better and her reputation drifts upwards (Corollary 5.2).

In the log-likelihood space, valuations are concave [Fig. 2(a)]. Concavity implies that incentives to reach for yield get stronger as reputation goes down [see Equation (14)]. Intuitively, as the reputation goes down, the probability of hitting the liquidation threshold becomes more sensitive to changes in reputation. Because liquidation is costly, the skilled expert becomes more risk-averse as she approaches liquidation.

Fig. 2(b) shows valuations in space of probabilities. The vertical line denotes the threshold where the opaque fund is liquidated. Investors pull out earlier than they pull out from a transparent fund, because they rationally expect the expert to reach for yield and to have lower expected returns. Because experts expect investors to pull out earlier, the investors' behavior feeds back into even stronger incentives to reach for yield.

Fig. 2(c) shows reaching-for-yield incentives $y(p)$. Reaching for yield goes down as reputation increases, especially for the skilled expert, who becomes less averse to reputation shocks as liquidation becomes less likely. This leads her to gradually shift from reaching for yield to "reaching for expected returns" as $y(p)$ goes to zero. Fig. 2(f) shows this directly by plotting the portfolio weight on the Tail portfolio $w_\kappa(p, \theta) = \frac{y(p)\lambda\kappa^+}{\sqrt{(\mu_S^+)^2 + (y(p)\lambda\kappa^+)^2}}$ as a function of reputation. As reputation grows, the skilled expert reduces her position on the Tail portfolio and increases her position on the Sharpe portfolio. This contrasts with the transparent benchmark, where the expert

[30] See Appendix E for a discussion of these parameter choices.

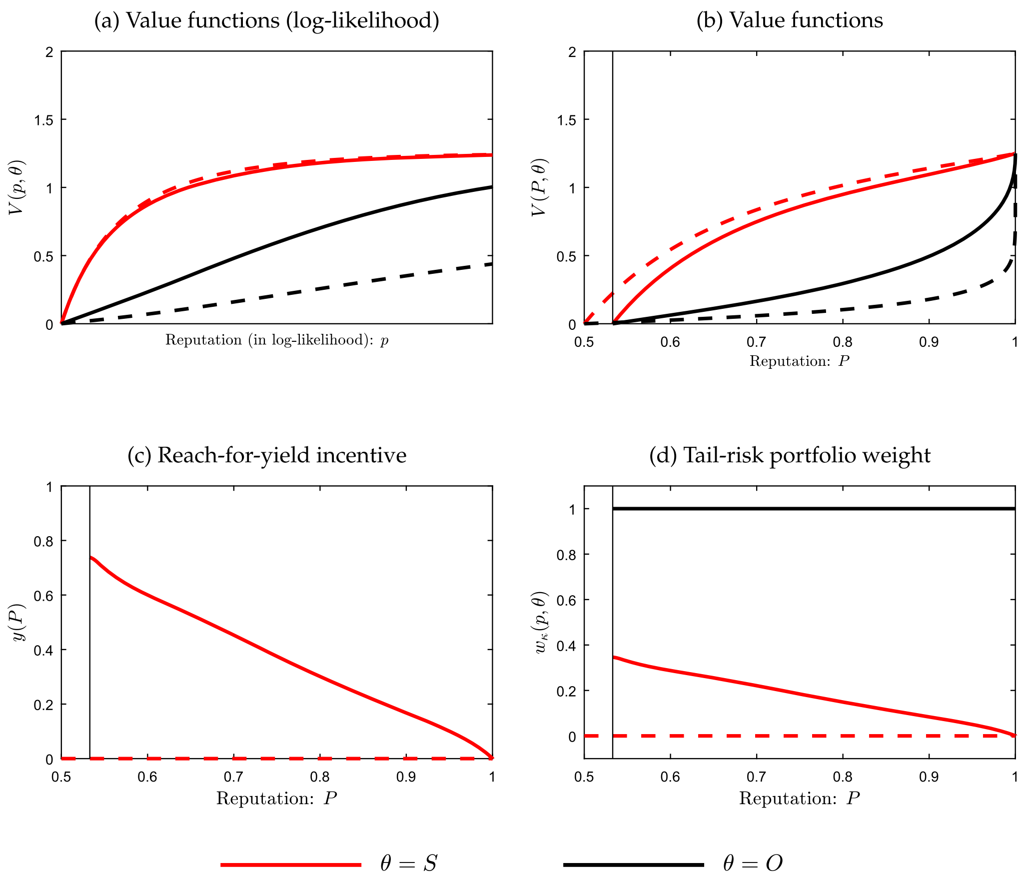

Fig. 2. **Valuation, incentives, and portfolio choice**. The top row shows valuations as a function of reputation. Panel (a) show valuations as function of reputation in log-likelihood space (p), and Panel (b) in probability space ($\mathbb{P}(p)$). The bottom row shows the policy functions. Panel (c) shows reach-for-yield incentives $y(p)$ as a function of reputation in probability space ($\mathbb{P}(p)$) and Panel (d) shows the resulting portfolio weight on the Tail portfolio $w_\kappa(p,\theta) = \frac{y(p)\lambda\kappa^+}{\sqrt{(\mu_S^+)^2+(y(p)\lambda\kappa^+)^2}}$. Black (red in the web version) lines denote the skilled expert and grey lines the opportunistic expert. The vertical bar denotes the threshold at which the opaque fund is liquidated. The transparent fund is liquidated in the left limit of the plots. Continuous lines denote the opaque fund and dashed lines the transparent fund.

always holds the Sharpe portfolio. The opportunistic expert, on the other hand, always invests on the Tail portfolio as a result of the strong incentives to reach for yield shown in Fig. 2(c).

Overall, the model predicts that the opportunist always reaches for yield more aggressively than the skilled expert, and the skilled expert's reaching-for-yield incentives peak as the fund approaches liquidation.

3.2. Capital allocation and expert performance

I decompose the fund returns in terms of the fund total volatility, which measures how much capital investors allocate to the expert, and the fund Sharpe Ratio, which is a measure of the quality of the expert capital allocation. Here I discuss the expert capital allocation decision. Fig. 3(a) shows that the skilled expert's yield increases as her reputation goes down. Heightened concerns with liquidation risk when reputation is low push the expert toward assets with high tail exposure because they over perform during normal times. Even though expected returns are constant across assets, reaching for yield reduces the diversification of the expert portfolio. The portfolio's Sharpe ratio falls as a result of this under diversification [Fig. 3(c)]. The opportunistic expert in this example is always constrained and will max out his tail exposure.[31] In the transparent benchmark, neither expert reaches for yield, and both simply hold the Sharpe portfolio.

Fig. 3(c) shows that the tail-event performance of the skilled expert improves with reputation, while the opportunist always performs equally poorly. Because reputation is a weighted average of past performance, this dynamic implies that bad performance is associated with higher yield and forecasts higher tail risk.

The dynamics of expected returns and liquidation have similarities to the mechanism proposed in Shleifer and Vishny (1997), where the risk of investors pulling out induces the manager to choose low-expected-return strategies. In their setting, trading strategies that are more volatile led to higher liquidation risk because investors were assumed not to adjust their behavior appropriately. A difference between their mechanism and mine is that here investor behavior is optimal. This difference is important because it allows me to study capital immobility.

3.3. *Capital flows*

For a given reputation, the total amount of capital allocated to a fund is a function of the fund's expected return; from Proposition 7, $\sigma_a(P_t) = \frac{P_t X(P_t, S)\mu_S - \psi_1}{2\psi_0}$. Fig. 4(a) shows that the opaque fund always attracts less capital and it is liquidated for a higher level of reputation. As reputation increases, and reach-for-yield incentives decreases, the capital allocated to an opaque fund converges to the transparent fund allocation.

The key determinant of the speed of capital flows is the rate at which investors learn. Figs. 3(c-e) depict the endogenous learning coefficients $g(p)$ and $g_J(p)$, and the rate of learning. The normal-times performance coefficient g decreases as reputation rises, while the tail-event coefficient increases. Intuitively, as the skilled expert reaches for yield less aggressively, the difference in normal-times performance across types shrinks. The reach for yield by the skilled type distorts allocations and decreases expected returns, but it *increases* the speed of learning. Thus, given the opportunistic type's behavior, there is a trade-off even from the vantage point of society as a whole. Transitory distortions in the allocation of capital lead to a faster convergence to the efficient allocation of capital. Portfolio opaqueness introduces a trade-off between the static and intertemporal efficiencies in the allocation of capital. Learning is faster when the skilled type reaches for yield aggressively, but gradually falls as her portfolio converges to the Sharpe portfolio. However, even when the learning rate is at its maximum, it is substantially lower than in the transparent benchmark.

Fig. 4(b) shows capital flows as a function of the capital gap, $\sigma_a(1) - \sigma_a(P)$, which is the difference between the long-run and the current level of capital. The slope of this plot measures the rate of convergence. A slope of 0.25 for the transparent fund case implies that the capital gap shrinks by 25% every year. Capital converges instead at 13% when the fund is opaque. Convergence rates are 50% slower than for the transparent fund.

[31] Recall that the opportunistic expert's expected returns are zero by assumption ($\mu_O = 0$).

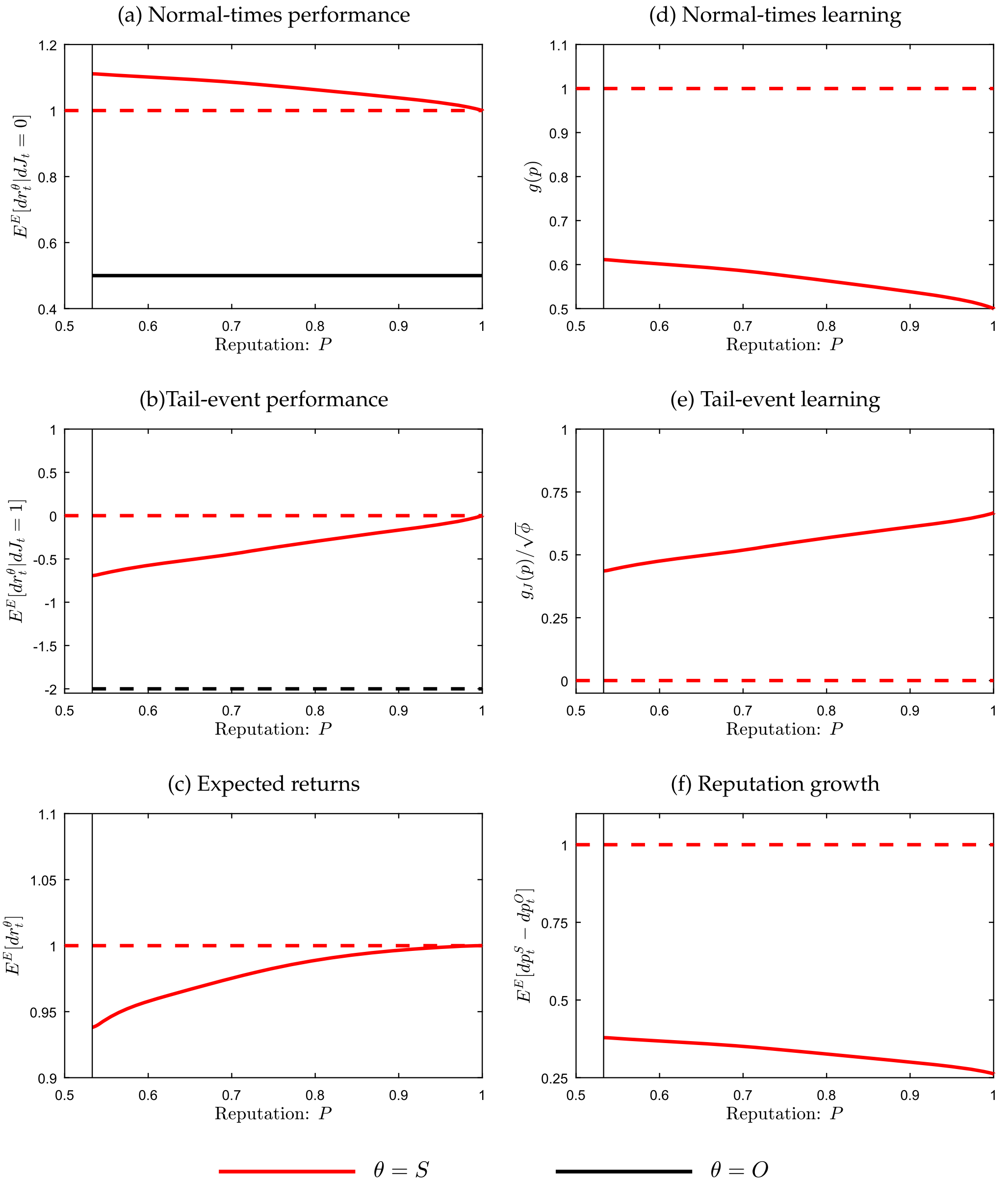

Fig. 3. **Performance and learning**. Panels (a), (b), and (c) show equilibrium performance during normal times, tail events, and fund expected returns. Panels (d) and (e) show the informativeness of performance during normal times and tail events. Panel (f) shows the expected growth of the skilled intermediary reputation. Black (red in the web version) lines denote the skilled expert and grey lines the opportunistic expert. The vertical bar denotes the threshold at which the opaque fund is liquidated. The transparent fund is liquidated in the left limit of the plots. Continuous lines denote the opaque fund and dashed lines the transparent fund.

Capital flows more slowly into the opaque fund because the opportunistic expert aggressively reaches for yield. This behavior reduces the information content of performance and slows down learning. The skilled expert strategic response only partially attenuates this reduction in the speed of capital, because reaching for yield is more costly for an expert with good investment opportunities.

3.4. The cross-sectional of capital flows

Here, I study how capital flows differently across technologies. The amount of capital in a technology—$X(P,S)\sigma_a(P)$—depends on how much capital a skilled expert has to invest—$\sigma_a(P)$—and how the skilled expert allocates this capital—$X(P,S)$.[32]

Fig. 4(c) shows that capital allocation to high-tail-risk technologies converges quickly to its long-run value and it even over shoots its long-run level. For technologies that appreciate in a tail event, we have exactly the opposite. Initially, these technologies experience capital outflows as investors allocate more capital to the fund. Experts short these technologies despite the fact that they would increase their expected returns. Experts with low reputation perceive the investment in these technologies as unattractive because they have low yields. They are unattractive when the expert has a short investment horizon due to the high liquidation risk.

Fig. 4(d) illustrates the speed of capital flows. We can see the capital flow as a function of the capital gap in a given technology. We see sharp differences across technologies when the fund is opaque. While capital flows into a high-tail-risk technology extremely quickly, capital initially flows out of the low-tail-risk technologies. In Figs. 4(c-d) we see that, when the fund is transparent, capital flows at the same rate to all technologies. Capital is equally scarce in each of them, so the expert allocates investments equally across them as more capital flows into the fund.

From an empiricist vantage point, the pattern in Fig. 4(d) is suggestive of market segmentation or investors' neglect of tail risk, but they are an equilibrium outcome in an environment where markets are integrated and investors are fully aware of the risks. Capital flows into perfectly integrated markets at very different rates. This pattern is driven by the endogenous investment incentives that arise from the way investors optimally allocate capital. Importantly, these distortions arise even though investors fully understand the environment and the risks they are exposed to. The lack of measurement leads to reaching-for-yield behavior and heterogeneity in the speed of capital flows.

3.5. Empirical implications

3.5.1. Flow sensitivity and reach-for-yield behavior

A key prediction of the model is that fund managers reach for yield more aggressively when investor flows are more sensitive to fund returns. Furthermore, the model also predicts that this relationship intensifies following outflows. Studying money market funds during the 2007-2009 financial crisis, Kacperczyk and Schnabl (2012) showed evidence consistent with these predictions. They showed that funds with flows that were more sensitive to performance were the ones that invested more in assets that paid out poorly as financial markets deteriorated in the fall of 2008, i.e., they took on more tail risk. They also showed that funds that suffered more severe

[32] I focus on the skilled expert allocation since all the technologies in the opportunist portfolio earn zero expected returns, and therefore, do not lead to any (direct) capital misallocation.

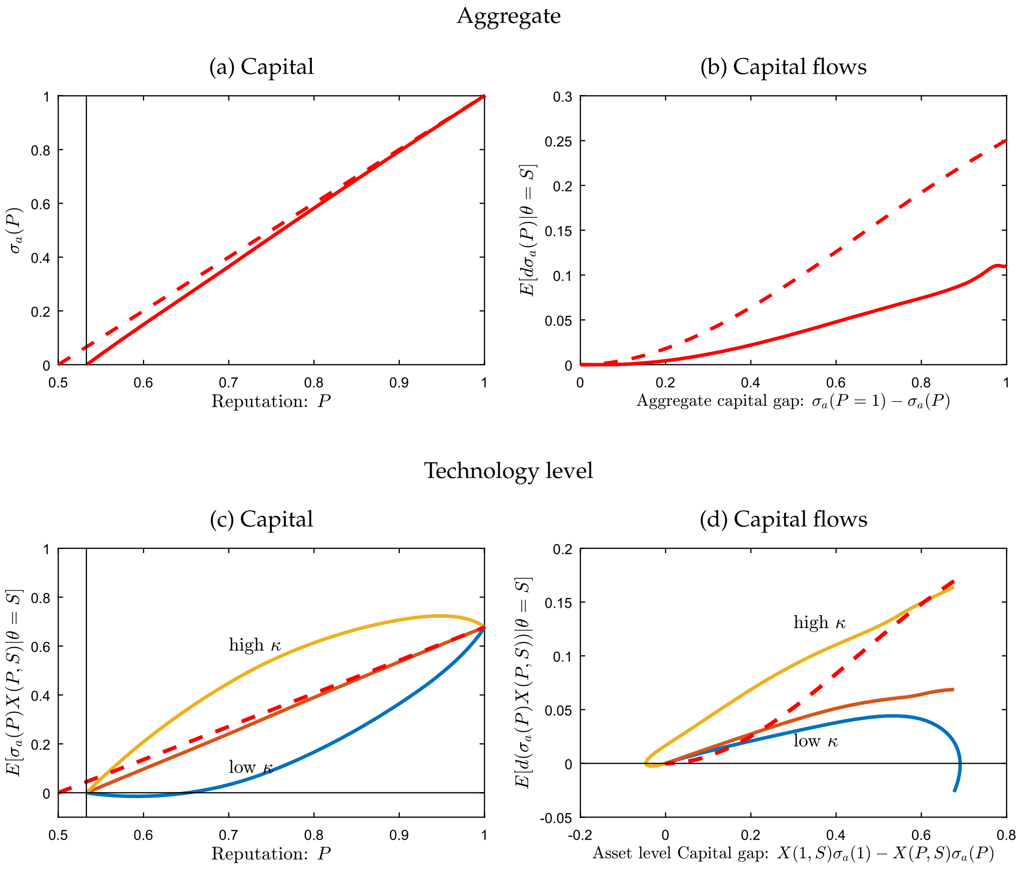

Fig. 4. **Capital immobility**. Panel (a) shows the total capital managed by the expert $\sigma_a(P_t)$ as a function of reputation. Panel (b) shows the expected aggregate inflow of capital for a skilled expert as a function of the capital gap ($\sigma_a(P = 1) - \sigma_a(P_t)$). Panel (c) shows the capital invested in three different technologies: with high tail risk (high κ), average tail risk, and low tail risk (low κ) as a function of reputation. Panel (d) shows the rate of capital inflow in these technologies as a function of the technology capital gap ($X(P = 1, S)\sigma_a(P = 1) - X(P_t, S)\sigma_a(P - t)$). The vertical bar denotes the threshold at which the opaque fund is liquidated. The transparent fund is liquidated in the left limit of the plots. Continuous lines denote the opaque fund and dashed lines the transparent fund.

outflows were also the funds that had more sensitive flows and that reached for yield more aggressively.

The academic literature and policy makers have interpreted this evidence as a result of the lack of market discipline: this pattern was driven either by investors' neglect of the link between higher fund yields and tail risk or because investors had confidence in a government bailout of these funds. In the model, it is exactly the market discipline that drives rampant reaching-for-yield behavior. Investors' response to fund manager incentives amplifies the incentives to reach for yield of investors and managers.

3.5.2. *Capital is less mobile in funds with more flexible mandates*

Greater investment mandate flexibility allows the expert to consider a wider set of assets for investment. An expansion in the investment opportunity set has two effects: higher alpha for the

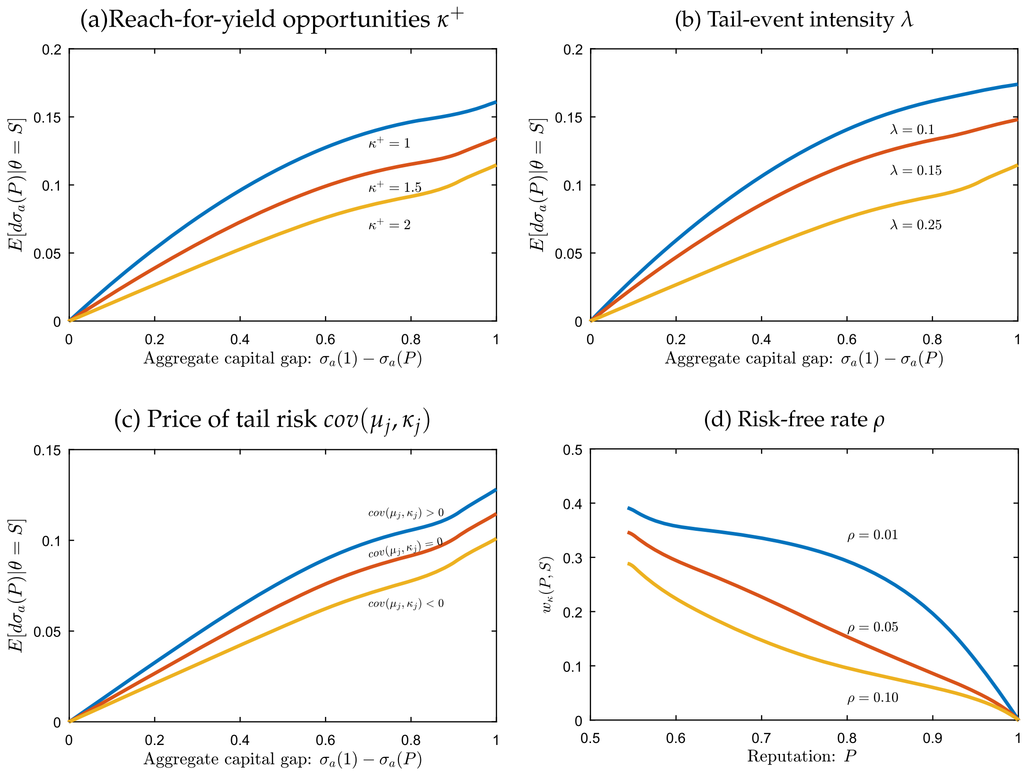

Fig. 5. **Empirical implications**. Panels (a) to (c) show the expected aggregate inflow of capital for a skilled expert as function of the capital gap for alternative parameters. Panel (a) varies κ^+, which denotes the maximum tail exposure feasible in the opportunity set. This shows how more flexible investment mandates reduce the speed of capital flows if they increase reach-for-yield opportunities. Panel(b) varies the tail-event intensity λ. This shows how an increase in the probability of a tail event can make capital flow more slowly because it increases reach-for-yield opportunities. Panel (c) varies the cross-sectional relationship between tail risk and expected returns—i.e., if tail-risk premium is positive or negative. This shows that capital will be particular slowmoving when tail risk is underpriced (negative tail-risk premium). Finally, panel (d) shows the portfolio weight of the skilled intermediary on the tail-risk portfolio $w_k(P, S)$ for different levels of the risk-free interest rate. This shows that reach-for-yield is stronger when monetary policy is looser.

skilled expert, but also greater opportunities to reach for yield for both types. Thus, the optimal mandate should be determined by balancing out these forces. Here I focus on the second effect, showing that an increase in the reach-for-yield opportunities, i.e., a higher κ^+, leads to more capital immobility.

Fig. 5(a) shows that the higher the κ^+, the more slowly capital flows. This result follows directly from the fact that the opportunistic type reaches for yield more aggressively than the skilled type. An increase in the performance that can be manufactured through tail risk reduces the (normal-times) performance difference across types, resulting in less learning and slower capital flows. A higher κ^+ also leads to more performance persistence and a more concave relation between flow and performance. Performance persistence is a by-product of slower learning and slower capital flows. The increasing sensitivity of flows after bad performance is the result of the stronger time variation in reach-for-yield behavior by the skilled type when κ^+ is higher.

An important source of variation in mandate flexibility is the asset class of a fund. The corporate bond market is an example of an asset class that exhibits low volatility, but substantial negative skewness (Bessembinder et al., 2009; Stein, 2013). This implies that capital should move particularly slowly into bond funds. It also suggests that, in contrast to equity mutual funds, bond funds should have performance that is more persistent, and the flow-performance relationship should be stronger after bad returns. There is recent evidence along these lines. Specifically, Goldstein et al. (2015) found that corporate bond fund flows were more sensitive to negative performance, in contrast to equity mutual funds, which tend to have a convex relationship between flows and performance.

3.5.3. Capital is less mobile when tail risk is underpriced

In the model, the trade-off between normal-times performance and expected returns emerges from the fact that the Tail portfolio and the Sharpe portfolio are different. This difference implies that the maximization of normal-times performance—that is, the reach for yield—reduces expected returns. The magnitude of the wedge between these portfolios depends critically on how expected returns are related to tail risk.

Thus far, I have assumed no relationship (a zero tail-risk premium). Fig. 5(c) shows the speed of capital flows as I vary the tail-risk premium. Capital moves more slowly when the tail-risk premium is negative. When tail risk is underpriced, there is a larger wedge between the portfolio that maximizes expected returns and the one that maximizes normal-times performance, because the Sharpe portfolio now involves shorting high-tail-risk technologies. This further attenuates the normal-times performance advantage of the skilled expert when they invest in the Sharpe portfolio. The end result is that normal-times performance is less informative and capital moves more slowly.

3.5.4. Capital is less mobile when a tail event is more likely

When tail risk is correctly priced, an increase in the probability of a tail event increases the normal-times performance that can be created by taking on tail risk. Fig. 5(b) shows the speed of capital flows for three values of λ, the tail-event intensity. The higher the tail intensity, the more slowly capital flows. Here why this happens: an increase in tail intensity increases the amount of performance that can be manufactured by reaching for yield; the difference in yields across types falls; performance becomes less informative; and the end result is slow-moving capital. I am not aware of any evidence that directly speaks to this prediction.

3.5.5. Reach for yield is stronger when interest rates are low

As noted by Rajan (2005, 2012), Stein (2013), and many others, sustained low interest rates are associated with strong reach-for-yield behavior in asset markets. Choi and Kronlund (2014) have documented evidence consistent with this view. In the model, the relation between reach-for-yield behavior and the level of interest rates emerges because incentives to reach for yield are tightly linked to the present value of future fees. Specifically, a lower interest rate increases the value of future fees, and consequently reduces the relative importance of performance incentives. This can be seen in Equation (14): a reduction in interest rates increases V_p but leaves m unchanged. As a result, reach for yield increases. To illustrate this relationship, Fig. 5(d) shows the skilled expert's portfolio weight on the Tail portfolio for three different levels of the risk-free rate. Consistent with the views discussed above, a lower interest rate leads to a higher portfolio weight on the Tail portfolio, i.e., stronger reach for yield.

3.5.6. Tail risk is underpriced in high-tail-risk assets

While the model in the main body of the text has no explicit implications for pricing, a simple extension where the expert trades against mean-variance investors, transforms the capital allocation predictions into asset pricing predictions (see Appendix B) . Intuitively, the very low allocation to the low-tail-risk assets translates into high expected returns, while the very high allocation to high-tail-risk assets translates into low expected returns.

Coval et al. (2009) documented evidence consistent with this prediction. Specifically, they showed that senior CDO tranches were persistently overpriced and junior CDO tranches persistently underpriced relative to option markets. Under the reasonable assumption that it is easier for investors to measure tail risks in an option portfolio than in a structured product portfolio, the model is consistent with both the underpricing of the junior tranches and overpricing of the senior tranches. Intermediaries dislike the junior tranche because it is very risky during normal times and has a low tail-risk-to-volatility ratio, while the senior tranche is close to safe during normal times and much more exposed to tail risk per unit of normal-times volatility, and therefore more appealing for a manager concerned with his or her short-term performance.

4. Conclusion

In this paper, I have developed a fully consistent narrative of how intermediaries and investors interact. I have shown how this interaction leads to time variation in reaching-for-yield incentives and has repercussions for the speed at which capital flows. Capital flows slowly to profitable opportunities. The reduction in capital flows is particularly severe for strategies that are good hedges against tail risks. The endogenous learning dynamic produces a feedback loop between liquidation risk and reach-for-yield behavior. In contrast with previous literature, the model results are driven by investors' sophisticated understanding of the environment. Information and capital flows shape and are shaped by the incentives intermediaries face.

The importance of reach-for-yield opportunities in slowing the flow of capital is the fundamental new insight of this framework. It points to a novel trade-off between the present and future allocation of capital in intermediated markets. Further, it complements the slow-moving capital literature by developing a mechanism that can account for a substantially more persistent misallocation of capital.

Appendix

This appendix contains: proofs and derivations used in the paper, a microfoundation for the cost function $c(\cdot)$, additional discussion of the importance of the model assumptions, discussion of the parameter choice, and the numerical method I used to solve the model.

Appendix A. Proofs

Proposition 1. *Investor learning about expert type.*

Proof. Consider the cumulative realized excess return history r_Δ in an interval Δ, and let P_t the perceived probability the intermediary is of type $\theta = S$ in the beginning of the interval. Bayes law implies,

$$P_{t+\Delta} = \frac{f(r_\Delta|\theta = S, P_t)P_t}{f(r_\Delta|\theta = S, P_t)P_t + f(r_\Delta|\theta = O, P_t)(1 - P_t)}, \tag{A.1}$$

where $f(r|\theta, P_0)$ is the probability distribution of a return history r if the intermediary is of type θ and initial reputation P_0. In our setting, this density is a complex object since the distribution of realized returns is time-varying due to the time-variation in the expert portfolio. However, in the dt limit, the problem simplifies because the portfolios become constant as the interval becomes arbitrary short. The problem is further simplified because in any interval dt, dJ_t equals to zero or one. The learning problem boils down to distinguish between two statistical models is two different observable states.

For this proof, I abstract form transaction costs just to economize on notation, but transaction costs have no impact on the learning dynamics since it impacts both experts equally.

First note that investors can detect whether $dJ_t = 1$ because the return quadratic variation perfectly reveals it. Let me start with normal times, $dJ_t = 0$. In this case, from the investor vantage point returns are distributed as,

$$r^\theta_{dt} \sim N(X^I(p,\theta)(\mu_\theta + \lambda\kappa)\,dt, 1dt). \tag{A.2}$$

Applying equation (A.1) I obtain

$$P_{t+dt} = \frac{N(r_{dt}|X^I(p,S)(\mu_S+\lambda\kappa)\,dt, dt)\times P_t}{N(r_{dt}|X^I(p,S)(\mu_S+\lambda\kappa)\,dt, dt)\times P_t + N(r_{dt}|X^I(p,O)(\mu_O+\lambda\kappa)\,dt, dt)\times(1-P_t)}. \tag{A.3}$$

Now lets go to log-likelihood space, define $p_{t+dt} = ln\left(\frac{P_{t+dt}}{1-P_{t+dt}}\right)$, and substitute to get,

$$\begin{aligned}
p_{t+dt} &= ln\Big(N(r_{dt}|X^I(p,S)(\mu_S+\lambda\kappa)dt, dt)\times P_t\Big)\\
&\quad - ln\Big(N(r_{dt}|X^I(p,O)(\mu_O+\lambda\kappa)\,dt, dt)\times(1-P_t)\Big)\\
&= p_t + ln\left(\exp\left(-\frac{\big(dr_t - X^I(p,S)(\mu_S+\lambda\kappa)\,dt\big)^2}{2dt}\right)\right)\\
&\quad - ln\left(\exp\left(-\frac{\big(dr_t - X^I(p,O)(\mu_O+\lambda\kappa)dt\big)^2}{2dt}\right)\right)\\
&= p_t - \frac{dr_t - X^I(p,S)(\mu_S+\lambda\kappa)\,dt}{2dt^2} + \frac{dr_t - X^I(p,O)(\mu_O+\lambda\kappa)\,dt}{2dt^2}\\
&= p_t + \Big(X^I(p,S)(\mu_S+\lambda\kappa) - X^I(p,O)(\mu_O+\lambda\kappa)\Big)\\
&\quad \times\left(dr_t - \frac{X^I(p,S)(\mu_S+\lambda\kappa) + X^I(p,O)(\mu_O+\lambda\kappa)}{2}dt\right).
\end{aligned} \tag{A.4}$$

Now let me define the normal-times learning coefficients as,

$$e(p_t) = E[dr^\theta_{t-}] = \frac{X^I(p,S)(\mu_S+\lambda\kappa) + X^I(p,O)(\mu_O+\lambda\kappa)}{2}dt, \tag{A.5}$$

$$g(p) = E[dr^S_{t-} - dr^O_{t-}] = \Big(X^I(p,S)(\mu_S+\lambda\kappa) - X^I(p,O)(\mu_O+\lambda\kappa)\Big), \tag{A.6}$$

to obtain the normal-times reputation dynamics

$$dp_{t-} = g(p)\,(dr_{t-} - e(p_t)). \tag{A.7}$$

Now lets focus on periods with tail events, $dJ_t = 1$. In this case we have $r_{dt}^{\theta} \sim N(-X^I(p,\theta)\kappa, \phi)$. Repeating exactly the same algebra as for the $dJ = 0$ case, and defining the tail-event learning coefficients as

$$e_J(p_t) = E[dr_t^{\theta}|dJ_t = 1] = \frac{X^I(p,S)\kappa + X^I(p,O)\kappa}{2}dt, \tag{A.8}$$

$$g_J(p) = E[dr_t^S - dr_t^O|dJ_t = 1] = -\left(X^I(p,S)\kappa - X^I(p,O)\kappa\right), \tag{A.9}$$

I obtain

$$p_{t+dt} = p_t + \frac{\left(-X^I(p,S)\kappa + X_I(p,O)\kappa\right)}{\phi}\left(r_{dt}^{\theta} + \frac{X^I(p,S)\kappa + X^I(p,O)\kappa}{2}\right), \tag{A.10}$$

$$dp_t = \frac{g_J(p)}{\phi}\left(r_{dt}^{\theta} - e_J(p_t)\right). \tag{A.11}$$

This proves Proposition 1. □

Corollary 1.1. *Reaching for yield and the speed of learning.*

Proof. Point (1) is immediate from Equation (8). If both hold the Sharpe portfolio we have $g(p) = \mu_S^+$ and $g_J(p) = 0$, then

$$\begin{aligned} &E[dp_t|\theta = S, X_S = X_\mu] - E[dp_t|\theta = O, X_O = X_\mu] \\ &= \mu_S^+\left(\mu_S^+ - \frac{\mu_S^+}{2}\right) - \mu_S^+\left(0 - \frac{\mu_S^+}{2}\right) && \text{(A.12)} \\ &= \frac{(\mu_S^+)^2}{2} - \frac{(\mu_S^+)^2}{2} && \text{(A.13)} \\ &= (\mu_S^+)^2 && \text{(A.14)} \end{aligned}$$

Point (2) also follows from Equation (8) by letting the skilled portfolio to be $X_S = X_\mu$ and the opportunist portfolio be defined by a tail exposure g_J. The expected reputation growth of each type is given by

$$\begin{aligned} &E[dp_t|\theta = S, X_S = X_mu] \\ &= (\mu_S^+ - g_J\lambda) \times \left(\mu_S^+ - \frac{\mu_S^+ + g_J\lambda}{2}\right) + \frac{g_J}{\phi} \times \left(0 + \frac{g_J}{2}\right)\lambda && \text{(A.15)} \\ &E[dp_t|\theta = O, (X_O - X_S)\kappa = g_J] \\ &= (\mu_S^+ - g_J\lambda) \times \left(0 - \frac{\mu_S^+ + g_J\lambda}{2}\right) + \frac{g_J}{\phi} \times \left(-g_J + \frac{g_J}{2}\right)\lambda && \text{(A.16)} \end{aligned}$$

Differentiating with respect to the tail exposure of the opportunist expert and evaluating it at $g_J = 0$ we show the result.

$$\frac{\partial E[dp_t|\theta = S, X_S = X_\mu] - E[dp_t|\theta = O, (X_O - X_S)\kappa = g_J]}{\partial g_J}|_{g_J=0} = -2\mu_S^+\lambda \le 0 \tag{A.17}$$

Point (3) also follows from Equation (8). Let the skilled portfolio be written as a function of a scalar x as $X_S(x) = F(x)X_* + xX_\kappa$, where $F(x) = \sqrt{1-x^2}$ implies that the portfolio $X_S(x)$ has variance equal to one for any x between 0 and 1. Let the opportunist portfolio be defined by a tail exposure κ^+. The expected reputation growth across types is given by

$$E[dp_t^S|X_S = X_S(x)] - E[dp_t^O|X_O\kappa = \kappa^+]$$
$$= (X_S(x)\mu_S - (\kappa^+ - X_S(x)\kappa)\lambda)^2 + \frac{(\kappa^+ - X_S(x)\kappa)^2}{\phi}\lambda. \tag{A.18}$$

Differentiating this reputation growth rate with respect to x, which is the tilt of the skilled type toward the Tail portfolio, I obtain

$$\frac{\partial E[dp_t|\theta = S, X_S = X_S(x)] - E[dp_t|\theta = O, X_O\kappa = \kappa^+]}{\partial x}|_{x=0} = \tag{A.19}$$
$$= 2(\mathcal{X}(0)\mu_S - (\kappa^+ - \mathcal{X}(0)\kappa)\lambda)\mathcal{X}'(0)(\mu_S + \lambda\kappa)2 - 2\frac{(\kappa^+ - \mathcal{X}(0)\kappa)}{\phi}\lambda\mathcal{X}'(0)\kappa$$
$$= 2(\mu_S^+ - \kappa^+\lambda)X_\kappa(\mu_S + \lambda\kappa) - 2\frac{\kappa^+}{\phi}\lambda X_\kappa\kappa$$
$$= 2(\mu_S^+ - \kappa^+\lambda)\lambda\kappa^+ - 2\frac{\kappa^+}{\phi}\lambda\kappa^+$$
$$= 2\lambda\kappa^+(\mu_S^+ - \kappa^+\lambda - \kappa^+\frac{1}{\phi}). \tag{A.20}$$

This implies that the reputation growth grows if an only if $\mu_S^+ - \kappa^+\lambda > \kappa^+/\phi$. This condition is fairly intuitive. It simply says that the signal-to-noise ratio of normal-times performance should be higher than the signal-to-noise ratio of tail risk performance at $x = 0$ (when there is no reach for yield) for reputation growth to be increasing in the skilled expert reach-for-yield. □

Proposition 2. *Learning speed when portfolios are transparent.*

Proof. There are two results here. The first result is that both skilled and opportunistic types invest in the Sharpe portfolio. That the skilled type invest in the Sharpe portfolio is immediate from her first order condition (Equation (11)). Given the skilled choice, the opportunistic must mimic her behavior because the portfolio is transparent and any tail risk taking would reveal the opportunistic type and lead to immediate liquidation. Given this first result, result 1 in Corollary 1.1 immediately imply the second result. □

Corollary 2.1. *Analytical solution for the transparent benchmark.* [*This Corollary is not in the main text.*]

Let's assume that fees are constant $a_t f_t = \overline{f}$, then experts' value functions are given by

$$V(p,\theta) = \frac{\overline{f}}{\rho}\left(1 - \exp\left(-\iota_\theta \times p\right)\right), \tag{A.21}$$

where $\iota_S = \frac{\sqrt{1+\frac{8\rho}{(\mu_S^+)^2}}+1}{2}$, $\iota_O = \frac{\sqrt{1+\frac{8\rho}{(\mu_S^+)^2}}-1}{2}$.

Proof. From Proposition 1, the optimal experts' portfolios are $X(p, S) = X(p, O) = X_*$.

This implies that $g(p_t) = \mu_S^+$. Substituting the optimal choice in the Bellman Equation (10), I obtain

$$\rho V(p, S) = a_t f_t + V_p(p, S)\frac{g^2}{2} + V_{pp}(p, S)\frac{g^2}{2}. \tag{A.22}$$

$$\rho V(p, O) = a_t f_t - V_p(p, O)\frac{g^2}{2} + V_{pp}(p, O)\frac{g^2}{2} \tag{A.23}$$

Using that $a_t f_t = \overline{f}$. Substituting in the ODE above I obtain the solution

$$V(p, S) = \frac{\overline{f}}{\rho} + K_1 e^{-\frac{1-\sqrt{1+\frac{8\rho}{g^2}}}{2}p} + K_2 e^{-\frac{1+\sqrt{1+\frac{8\rho}{g^2}}}{2}p} \tag{A.24}$$

$$V(p, O) = \frac{\overline{f}}{\rho} + K_3 e^{-\frac{-1-\sqrt{1+\frac{8\rho}{g^2}}}{2}p} + K_4 e^{-\frac{-1+\sqrt{1+\frac{8\rho}{g^2}}}{2}p}. \tag{A.25}$$

Imposing the boundary conditions $V(0, \theta) = 0$ and that $V(\infty, \theta) = \frac{\overline{f}}{\rho}$ I obtain the result. □

Lemma 1. *Scalar representation of the skilled portfolio.*

Proof. From the first order condition in Equation (12) we immediately have that the optimal skilled expert portfolio is a combination of $\mu_S \Sigma^{-1}$ and $\kappa \Sigma^{-1}$. Together with the unit-variance constraint, it is immediate that we can represent the optimal portfolio as $X_* x_1 + X_\kappa x_2$ where $\sqrt{x_1^2 + x_2^2} = 1$. The representation in this Lemma is a particular case of this representation. The scalar y can be interpreted as reach-for-yield incentives because if $y = 0$, the portfolio is the Sharpe portfolio, the portfolio with maximum expected returns, and if $y = 1$, the portfolios is the combination that maximizes the fund yield, i.e., normal-times performance. See the proof of Proposition 3 below for additional detail. □

Proposition 3. *Skilled expert portfolio choice.*

Proof. I start by deriving the first order condition in Equation (12), and then show how the optimal choice can be characterized by the incentive schedule that satisfies the fixed point equation (14). Let me start by isolating the terms in the Bellman equation (Equation (10)) that depend directly on the expert's choice,

$$\begin{aligned} &\sup_{X \in \Omega} V_p g(p_t) X (\mu_\theta + \lambda\kappa) \\ &\quad + \lambda \mathbb{E}_t^E \left[V\left(p + g_J(p) \times \left(-X\kappa - \mathbb{E}^I\left[-X^I \kappa | p\right]\right) + g_J(p)\epsilon, \theta\right) | dJ_t = 1 \right]. \end{aligned} \tag{A.26}$$

I conjecture for now that the value function is positive, increasing and concave in reputation ($V(p, \theta) \geq 0$, $V_p(p, \theta) \geq 0$, $V_{pp} < 0$) and positive return surprises are always good news about expert type ($g(p), g_J(p) \geq 0$). I will show later that four of these properties always hold in equilibrium. I show that concavity holds in specific cases and verify that holds numerically more generally.

It follows from these properties that the third term is decreasing in the portfolio tail-exposure $X\kappa$, while the other terms are linear in the portfolio expected return and normal-times performance. Note that

$$\kappa\kappa^\top g_J(p)^2\lambda\mathbb{E}_t^E \\ \times\left[V_{pp}\left(p+g_J(p)\times\left(-X\kappa-\mathbb{E}^I\left[-X^I\kappa|p\right]\right)+g_J(p)\epsilon,\theta\right)|dJ_t=1\right]<0, \tag{A.27}$$

so the first order condition is necessary and sufficient to characterize the optimal policy.

Differentiating with respect to X, I obtain Equation (12). Solving for X, I get

$$X=\zeta^{-1}\left(\mu_S\times y_1+y_2\times\lambda\times\kappa\right)^\top\Sigma^{-1}, \tag{A.28}$$

where y_1 and y_2 are scalars given by

$$y_1=V_p(p,\theta)g(p) \tag{A.29}$$

$$y_2=V_p(p,S)g(p)-g_J(p_t)\mathbb{E}_\epsilon^E \\ \times\left[V_p\left(p+g_J(p)\times\left(-X\kappa-\mathbb{E}^I\left[-X^I\kappa|p\right]\right)+g_J(p)\epsilon,S\right)\right]$$

Substituting in the unit-variance constraint, $X\Sigma X^\top=1$, we get,

$$\zeta=\sqrt{\left(\mu_S\times y_1+y_2\times\kappa\right)^\top\Sigma^{-1}\left(\mu_S\times y_1+y_2\times\kappa\right)}. \tag{A.30}$$

Define $y=y_2/y_1$ and using that $\mu_S'\Sigma\kappa=0$, $X_*\propto\mu_S\Sigma^{-1}$, and $X_\kappa\propto\kappa\Sigma^{-1}$, we can write

$$X=X_*\frac{\mu_S^+}{\sqrt{(\mu_S^+)^2+(y\lambda\kappa^+)^2}}+\frac{y\lambda\kappa^+}{\sqrt{(\mu_S^+)^2+(y\lambda\kappa^+)^2}}X_\kappa. \tag{A.31}$$

Now note that y depends on the choice $X(p,S)$. So the above equation characterizes X only implicitly. Because X is of the same dimension as the number of assets, this fixed point problem in Equation (A.31) is a hard one to solve. I will now characterize the solution in terms of scalar incentives y. Define $\mathcal{X}(y)$ as the function that maps an incentive y into a portfolio choice X,

$$\mathcal{X}(y)=X_*\frac{\mu_S^+}{\sqrt{(\mu_S^+)^2+(y\lambda\kappa^+)^2}}+\frac{y\lambda\kappa^+}{\sqrt{(\mu_S^+)^2+(y\lambda\kappa^+)^2}}X_\kappa. \tag{A.32}$$

Now equilibrium is much simpler. It must be the case that a given incentive y imply a choice $\mathcal{X}(y)$, which itself is consistent with incentive y. Formally for any p, y must solve

$$y=\frac{V_p(p,S)g(p)-g_J(p_t)/\phi\mathbb{E}_\epsilon^E\left[V_p\left(p+g_J(p)/\phi\left(-\mathcal{X}(y)\kappa-\mathbb{E}^I\left[-X^I\kappa|p\right]\right)+g_J(p)/\sqrt{\phi}\epsilon,S\right)\right]}{V_p(p,S)g(p)}, \tag{A.33}$$

Instead of solving for the equilibrium portfolio that is consistent with equilibrium, we solve for the incentive schedule that is consistent with equilibrium. For a given reputation that is simply an scalar. This proves Proposition 3. □

Proposition 4. *Opportunist portfolio.*

Proof. I start from the first order condition in Equation (12),

$$X_t=\zeta_t^{-1}\left(g(p_t)\lambda V_p-\lambda g_J(p_t)/\phi\mathbb{E}^E\left[V_p\left(p_{t+},O\right)\right]\right)\Sigma^{-1}\kappa, \tag{A.34}$$

where $\mu_O=0$ implies the Sharpe ratio maximization incentive m drops out. It is immediate that if the unit-variance constraint binds $\zeta_t>0$ we have that $X_t=X_\kappa$ or $X_t=-X_\kappa$. This happens if $|\left(g(p_t)\lambda V_p - \lambda g_J(p_t)/\phi \mathbb{E}^E\left[V_p\left(p_{t+}, O\right)\right]\right)| > 0$, i.e., if reach-for-yield incentives are not zero in equilibrium. Because in this case only the sign of the incentives matter, we can write reach-for-yield incentives as in Equation (17). When reach-for-yield incentives are equal to zero, then the opportunistic expert is indifferent between any portfolio and the equilibrium portfolio is the one consistent with a zero reach-for-yield incentive. That is, g_J is pinned down by Equation (17). When reach-for-yield incentives are positive we have that $X_t = X_\kappa$ and $g_J = \kappa^+ - \mathcal{X}(y^I)\kappa$. When they are negative we have the other extreme and $X_t = -X_\kappa$ and $g_J = -\kappa^+ - \mathcal{X}(y^I)\kappa$. I prove in Corollary 5.2 below that this case will never happen in equilibrium. □

Proposition 5. *Equilibrium portfolios.*

Proof. Equation (18) comes from Proposition 3 and it is as the first order condition of the skilled type together with the equilibrium restriction that investors beliefs are consistent with the skilled expert policy. The second equation comes from Proposition 4 and can be thought as the first order condition of the opportunistic type with the added restriction that $\mathcal{D}_O = 0$ only holds when the constraint $(g_J + \mathcal{X}(\pi)\kappa) \leq \kappa^+$ does not bind.

The third equation comes from the fact that $\mathcal{D}_O$ cannot be negative in Equilibrium. Note that if $\mathcal{D}_O < 0$, it follows that $g_J < 0$, but from Equation (17) that implies $\mathcal{D}_O > 0$. Therefore $\mathcal{D}_O > 0$ never happens in equilibrium. The last inequality simply state that the tail risk of the opportunist expert portfolio is bounded by the investment opportunity set. □

Corollary 5.1. *The skilled expert always reaches for yield.*

Proof. Note that concavity implies $V_{pp} < 0$. Together with $\partial \mathcal{X}(y)\mu/\partial y < 0$, this implies $\partial \mathcal{D}_S(p, y|g_J^I, y)/\partial y < 0$ if $g_J^I > 0$. So because y is increasing in y and $\mathcal{D}_S(p, y|g_J^I, y)$ is decreasing in y, then they must cross at most once. Note that if $\mathcal{D}_S(p, 0|g_J^I, 0) > 0$ and $\mathcal{D}_S(p, 1|g_J^I, 1) < 1$, they would cross at least once.

Lets start by showing that $\mathcal{D}_S(p, 1|g_J, 1) < 1$. Note that $g_J \geq 0$ and $V_p \geq 0$. So if $\mathcal{X}(y = 1)\mu_S - \lambda g_J \geq 0$, the result follows. Lets assume g_J is maximum at $g_J = \kappa^+ - \mathcal{X}(y = 1)\kappa$, then $\mathcal{X}(y = 1)(\mu_S + \lambda\kappa) - \lambda\kappa^+ = \sqrt{(\mu_S^+)^2 + (\lambda\kappa^+)^2} - \lambda\kappa^+ > 0$.

To prove the result remains to show $\mathcal{D}_S(p, y = 0|g_J^I, 0) > 0$. Note that

$$
\begin{aligned}
&Sign[\mathcal{D}_S(p, y = 0|g_J, 0)] \\
&= Sign\left[(\mu_S^+ - \lambda g_J) - \lambda\frac{g_J}{\phi}\frac{\mathbb{E}^E\left[V_p\left(p + \frac{g_J}{\phi} \times \left(\frac{g_J}{2}\right) + \frac{g_J}{\sqrt{\phi}}\widetilde{Y}, S\right)\right]}{V_p(p, S)}\right],
\end{aligned}
\tag{A.35}
$$

and that the ratio $\frac{\mathbb{E}^E\left[V_p\left(p + (\frac{g_J}{\phi})^2/2 + \frac{g_J}{\sqrt{\phi}}\widetilde{Y}, S\right)\right]}{V_p(p,S)}$ is bounded since $0 \leq g_J \leq \kappa^+$ and V is continuous and bounded. For example, for the transparent benchmark (Corollary 2.1 in this Appendix) this ratio is simply

$$
\exp\left(-\iota_S(1 - \iota_S)(\frac{g_J}{\sqrt{\phi}})^2/2)\right),
\tag{A.36}
$$

where $\iota_S = \frac{\sqrt{1+\frac{8\rho}{(\mu_S^+)^2}}+1}{2}$. Therefore, for ϕ that is large enough $\mathcal{D}_S(p,0|g_J,0) > 0$, since $(\mu_S^+ - \lambda g_J) > 0$ for any $g_J \leq \kappa^+$. This proves that indeed $y > 0$ and the skilled always reaches for yield. □

Corollary 5.2. *Reputation is increasing in performance.*

Proof. First note that if $\mathcal{D}_O < 0$, $g_J < 0$, but that implies $\mathcal{D}_O > 0$, therefore it will never hold in equilibrium. It follows that $g_J \geq 0$. If $g_J = 0$, then $\mathcal{D}_O = 1$, what implies $g_J > 0$, so it can't be an equilibrium either. It follows that $g_J > 0$. This proves that reputation is always increasing in tail-event performance.

It remains to prove that $g = \mathcal{X}(y)\mu_S - \lambda g_J \geq 0$. Suppose $\mathcal{X}(y)\mu_S - \lambda g_J < 0$. Then it follows from (17) that $\mathcal{D}_O < 0$. Proposition 4 implies $g_J < 0$ and that $\mathcal{X}(y)\mu_S - \lambda g_J > 0$ for any $y \in [0,1]$. This is inconsistent with an equilibrium where $g(p) < 0$. From Corollary 5.1 we have that $y \in [0,1]$. This proves the result. □

Corollary 5.3. *Learning is slower when portfolios are opaque.*

Proof. From Equation (8), we have that the speed of learning is

$$\mathbb{E}[|dp_t||\theta] = \frac{(X(p,S)\mu_S)^2}{2}\left[\left(1 - \frac{g_J(p)\lambda}{X(p,S)\mu_S}\right)^2 + \frac{1}{\lambda\phi}\left(\frac{g_J(p)\lambda}{X(p,S)\mu_S}\right)^2\right]. \tag{A.37}$$

From Corollaries 5.2 we have that $g_J(p) \geq 0$ and $g(p) \geq 0$. This implies that $\frac{g_J(p)\lambda}{X(p,S)\mu_S} \in [0,1]$. If $\phi > 1/\lambda$, then $\left[\left(1 - \frac{g_J(p)\lambda}{X(p,S)\mu_S}\right)^2 + \frac{1}{\lambda\phi}\left(\frac{g_J(p)\lambda}{X(p,S)\mu_S}\right)^2\right]$ is bounded above by 1. It follows that $\mathbb{E}[|dp_t||\theta] \leq \frac{(X(p,S)\mu_S)^2}{2}$. □

Lemma 2. *Equilibrium existence.*

Proof. Suppose ϕ is sufficiently high such that $\mathcal{D}_O > 0$ and $g_J = \kappa^+ - \mathcal{X}(y)\kappa$ for any y. Under the conditions of Corollary 5.1 we have that $\mathcal{D}_S(p, y=1|g_J^I) < 1$ and $\mathcal{D}_S(p,0|g_J^I) > 0$ for any $g_J^I > 0$. Since $\kappa^+ - \mathcal{X}(y)\kappa > 0$, it applies here. Because $\mathcal{D}_S(p,y)$ and y are continuous, they must cross at least once. This proves that for high enough ϕ there always exist an equilibrium. □

Definition 1. Equilibrium selection: defining the fastest equilibrium. For any $p \geq 0$, let $\mathcal{M}(p) = \{\{y, g_J\} \in [0,1] \times [0,\kappa^+] | \{y, g_J\}$ satisfies Proposition 5$\}$ be the set of incentives consistent with equilibria for a given reputation, then the fastest equilibrium satisfies

$$y^*(p,S) = \arg \max_{\{y,g_J\}\in\mathcal{M}(p)} (\alpha_s \mathcal{X}(y)\mu - \lambda g_J)^2 + \lambda\left(\frac{g_J}{\sqrt{\phi}}\right)^2. \tag{A.38}$$

Proposition 6. *Expert liquidation.*

Proof. Investors liquidate the expert when they can no longer break even for any positive fund size, that is when $\mathbb{E}_t^I[\sigma_{r,t} dr_t^\theta - f dt] < 0$ for any positive investment a_t in the fund. Plugging Equation (24), I obtain

$$\sigma_r(\underline{P})\underline{P}X_I(\underline{P},S)\mu_S - \sigma_r(\underline{P})c(0) - f = 0.$$

This proves the proposition if you substitute $f = 0$. □

Proposition 7. *Equilibrium supply of capital.*

Proof. Plugging equation (24) into $a_t f_t$ I obtain

$$a_t\sigma_{r,t}\underline{P}X_I(\underline{P},S)\mu_S - a_t\sigma_{r,t}c(a_t\sigma_{r,t}). \tag{A.39}$$

Differentiating with respect to $a_t\sigma_{r,t}$, Equation (26) follows. Substituting the transaction cost function $c(\cdot)$, I can solve for the optimal fund size $\sigma_{a,t} = a_t\sigma_{r,t}$ and total dollar fees $a_t f_t$.

In the case of a quadratic cost function I obtain

$$\sigma_{a,t}\psi_1 = PX(P,S)\mu_S - \left(\psi_0 + \psi_1\sigma_{a,t}\right), \tag{A.40}$$

which yields $\sigma_{a,t} = (PX(P,S)\mu_S - \psi_0)/(2\psi_1)$ as enunciated in this Proposition. To obtain the optimal fee that implements the optimal size choice, I substitute in the investors break-even condition [Equation (24)]. □

Proposition 8. *Capital flows with opaque and transparent funds.*

Proof. From Proposition 6 we have that $\sigma_a(p, y=0) = (\mathbb{P}(p)\mu_S^+ - \psi_0)/\psi_1$. We have from Corollary 1.1 that $E[dp|y=0,\theta=S] = (\mu_S^+)^2/2$ and $var[dp|y=0] = (\mu_S^+)^2$. It follows from Ito's lemma that

$$E[d(\sigma_a^S(p,y=0)] = \mathbb{P}'(p)\frac{\mu_S^+}{\psi_1}(\mu_S^+)^2 + \mathbb{P}''(p)\frac{\mu_S^+}{\psi_1}(\mu_S^+)^2 \tag{A.41}$$

$$E[d(\sigma_a^O(p,y=0)] = -\mathbb{P}'(p)\frac{\mu_S^+}{\psi_1}(\mu_S^+)^2 + \mathbb{P}''(p)\frac{\mu_S^+}{\psi_1}(\mu_S^+)^2. \tag{A.42}$$

Equation (27) follows.

Now for the opaque portfolio case we have,

$$\begin{aligned}E[d(\sigma_a^S(p,y)] = {} & \mathbb{P}'(p)\frac{\mathcal{X}(y)\mu_S}{2\psi_1}(\mathcal{X}(y)\mu_S - g_J(y)\lambda)^2\\ & + \mathbb{P}''(p)\frac{\mu_S^+}{2\psi_1}(\mu_S^+)^2\\ & + \lambda E[(\mathbb{P}(p_+) - \mathbb{P}(p))\mathcal{X}(y)\mu_S/\phi_1|\theta = S]\end{aligned} \tag{A.43}$$

$$\begin{aligned}E[d(\sigma_a^O(p,y)] = {} & -\mathbb{P}'(p)\frac{\mathcal{X}(y)\mu_S}{2\psi_1}(\mathcal{X}(y)\mu_S - g_J(y)\lambda)^2\\ & + \mathbb{P}''(p)\frac{\mu_S^+}{2\psi_1}(\mu_S^+)^2\\ & + \lambda E[(\mathbb{P}(p_+) - \mathbb{P}(p))\mathcal{X}(y)\mu_S/\phi_1|\theta = O]\end{aligned} \tag{A.44}$$

$$\begin{aligned}E[d(\sigma_a^S(p,y) - d\sigma_a^O(p,y))] = {} & \mathbb{P}'(p)\frac{\mathcal{X}(y)\mu_S}{\psi_1}(\mathcal{X}(y)\mu_S - g_J(\pi)\lambda)^2 +\\ & + \lambda\left(E[(\mathbb{P}(p_+)|\theta = S]\right.\\ & \left.-E[(\mathbb{P}(p_+)|\theta = O]\right)\mathcal{X}(y)\mu_S/\phi_1\end{aligned} \tag{A.45}$$

A second-order Taylor approximation for the difference in after tail event reputations yields (the second order terms wash out because they are the same for both types),

$$(E[(\mathbb{P}(p_+)|\theta = S] - E[(\mathbb{P}(p_+)|\theta = O]) \approx \mathbb{P}'(p)\frac{g_J(y)^2}{\phi}.$$

Plugging back in the Equation (A.45) above I obtain Equation (28). □

Corollary 8.1. *Capital flows slower when the fund is opaque.*

Proof. The proof is an immediate implication of taking the limit $\phi \to \infty$ for Equation (28) together with the fact that $g_J(y) > 0$. □

Corollary 8.2. *Capital flows faster when the skilled expert reaches for yield.*

Proof. Differencing out Equation (28) with respect to y we obtain

$$\begin{aligned}
&\frac{\partial E\left[d(\sigma_a^S(p,y) - \sigma_a^O(p,y))|y\right]}{\partial y} \\
&\quad = \frac{\mathbb{P}'(p)\mathcal{X}(y)\mu_S}{\psi_1} 2(\mathcal{X}(y)\mu_S - g_J(y)\lambda)(\mathcal{X}'(y)\mu_S - g_J'(y)\lambda). \\
&\qquad + \frac{\mathbb{P}'(p)\mathcal{X}(y)\mu_S}{\psi_1} 2\left(\lambda\frac{g_J(y)g_J'(y)}{\phi}\right) \\
&\qquad + \frac{\mathbb{P}(p)\mathcal{X}'(y)\mu_S}{\psi_1}\left(\mathcal{X}(y)\mu_S - g_J(y)\lambda)^2 + \lambda\frac{g_J^2(y)}{\phi}\right).
\end{aligned} \tag{A.46}$$

Recognizing that $\mathcal{X}'(y=0) = \frac{\lambda\kappa^+}{\mu_S^+}X_\kappa$ and $\mathcal{X}(y=0) = X_*$, it follows that

$$\frac{\partial E\left[d(\sigma_a^S(p,y) - \sigma_a^O(p,y))|y\right]}{\partial y}\Big|_{y=0} = \frac{\mathbb{P}'(p)\mu_S^+}{\psi_1} 2\left(\mu_S^+ - \kappa^+(\lambda + \frac{1}{\phi})\right)\frac{(\lambda\kappa^+)^2}{\mu_S^+}, \tag{A.47}$$

and from this it is immediate that

$$Sign\left[\frac{\partial E\left[d(\sigma_a^S(p,y) - \sigma_a^O(p,y))|y\right]}{\partial y}\Big|_{y=0}\right] = Sign\left[\mu_S^+ - \kappa^+(\lambda + \frac{1}{\phi})\right]. \tag{A.48}$$

This proves the result. □

Proposition 9. *Equilibrium valuations: an analytical solution.*

Proof. From Proposition 3 it is immediate that $\phi \to \infty$ implies $y = 1$. Therefore the optimal portfolio for the skilled expert is $\mathcal{X}(y=1) = \frac{\mu_S^+}{\sqrt{(\mu_S^+)^2 + (\pi\lambda\kappa^+)^2}}X_* + \frac{\lambda\kappa^+}{\sqrt{(\mu_S^+)^2 + (\lambda\kappa^+)^2}}X_\kappa$ and from Proposition 4 the opportunist is in a corner, choosing κ^+.

This implies that $g(p_t) = \sqrt{(\mu_S^+)^2 + (\lambda\kappa^+)^2} - \kappa^+$. Substituting in the Bellman Equation (10) the optimal choice, I obtain

$$\rho V(p,S) = a_t f_t + V_p(p,S)\frac{g^2}{2} + V_{pp}(p,S)\frac{g^2}{2}. \tag{A.49}$$

$$\rho V(p,O) = a_t f_t - V_p(p,O)\frac{g^2}{2} + V_{pp}(p,O)\frac{g^2}{2} \tag{A.50}$$

Substituting $a_t f_t = \overline{f}$ in the ODE above, I obtain

$$V(p,S) = \frac{\overline{f}}{\rho} + K_1 e^{-\frac{1-\sqrt{1+\frac{8\rho}{g^2}}}{2}p} + K_2 e^{-\frac{1+\sqrt{1+\frac{8\rho}{g^2}}}{2}p} \tag{A.51}$$

$$V(p,O) = \frac{\overline{f}}{\rho} + K_3 e^{-\frac{-1-\sqrt{1+\frac{8\rho}{g^2}}}{2}p} + K_4 e^{-\frac{-1+\sqrt{1+\frac{8\rho}{g^2}}}{2}p}. \tag{A.52}$$

Imposing the boundary conditions $V(0,\theta) = 0$ and that $V(\infty,\theta) = \frac{\overline{f}}{\rho}$, I obtain the solution. □

Appendix B. Microfoundations for the transaction cost function $c(\cdot)$

We now provide a microfoundation for the source of decreasing returns to scale for an expert. Specifically, we provide a microfoundation for the cost function $c(\cdot)$. This section also provides a natural motivation for the fact that experts are exposed to idiosyncratic shocks and face different investment opportunities, specifically a different vector of expected returns μ_θ. The idea is that each expert has knowledge and trades in a particular local market where local investors have hedging needs with respect to market-specific risks. Together these ingredients generate heterogeneity in skill, decreasing returns to investment at the expert level, and idiosyncratic risk that are idiosyncratic to the expert.

An expert of type θ trades n assets with a representative local hedger. The local hedger has mean-variance preferences with risk-aversion ψ_1 and endowment vector E_θ. The vector of endowments controls the gains from trade between the expert and the local hedger. A skilled expert has access to a market where local hedgers have large hedging demands $|E_S| > 0$, while the opportunistic type trades in a market where local hedgers do not have any hedging needs $E_O = 0$. Regardless of their type, the expert has operational costs that scale with the riskness of the fund position $\psi_0\sigma_a$. While it would be puzzling that managers would have to pay a fixed cost simply to hold an asset, these assets should be thought more broadly as investment strategies which require costly trading. For example, holding a portfolio that buys the market for high-yield bonds involves substantial churning and transaction costs.

The utility of the local hedger if they sell vector Q of local assets is

$$U((E_\theta - Q)dR) = E[(E_\theta - Q)(dR - \rho)] - \frac{\psi_1}{2} Var((E_\theta - Q)dR), \tag{B.53}$$

which yields the following first order condition: $E[dR - \rho] = \psi_1 Var(dR)(E_\theta - Q)$. This FOC pins down expected returns of the local assets, given an expert asset allocation Q. Using the decomposition $Q = aX\sigma_r$, where X is a unit-variance position, σ_r is the fund return volatility, and a is the fund assets under management, we can write the expert total dollar returns as

$$\begin{aligned} QE[dR - \rho] - \psi_0\sigma_a &= Q\psi_1 Var(dR)(E_\theta - Q) - \psi_0\sigma_a && \text{(B.54)} \\ &= \sigma_a X\psi_1 Var(dR)(E_\theta - \sigma_s X) - \psi_0\sigma_a && \text{(B.55)} \\ &= \sigma_a\psi_1 X Var(dR)E_\theta - \psi_1\sigma_a^2 - \psi_0\sigma_a && \text{(B.56)} \end{aligned}$$

$$\frac{QE[dR]}{\sigma_a} = X\psi_1 Var(dR)E_\theta - \psi_1\sigma_a - \psi_0 \tag{B.57}$$

We now define $\mu_\theta = \psi_1 Var(dR)E_\theta$ and $c(\sigma_a) = \psi_0\sigma_a + \psi_1\sigma_a^2$ to obtain the fund expected results in the following form:

$$QE[dR] = \sigma_a X\mu_\theta - c(\theta). \tag{B.58}$$

This is exactly the functional form we adopt through out the paper.

Note that operational costs ψ_0 are important only to guarantee that an expert is liquidated for a reputation $P > 0$. This creates liquidation risk and is essential for the paper mechanism. The assumption that operational costs increase with asset risk σ_a is not important. Operational costs could also be of the fixed cost type and the cost function given by $c(\sigma_a) = \psi_0 + \psi_1\sigma_a^2$. The only important difference is that the expert would be liquidated before the fund assets reached zero due to the fixed operational costs.

We see that the opportunistic expert has zero expected returns because $E_O = 0$, but as he invests more assets in the local market, he must invest at increasingly more negative returns to compensate local hedgers for the risk that they must bear. The skilled type earns expected return $\mu_S > 0$ in the first unit invested, because local hedgers are exposed to risk due to the asset endowment that they own. As local hedgers sell more of their endowment, the required compensation for bearing risk goes down, leading to a reduction in expected returns.

This framework also directly links the model's quantity predictions to predictions about equilibrium expected returns. Specifically, assets that receive larger flows have lower expected returns in equilibrium. This can immediately be seen from the local hedgers first-order condition: $E[dR - \rho] = \psi_1 Var(dR)(E_\theta - Q)$, i.e., an increase in Q_i reduces asset i expected returns.

Appendix C. Numerical solution

I apply the finite-difference method to solve the integro-differential equation (10). To solve for optimal policies, I sequentially iterate until the value functions converges. The pair of value functions $\{V(p, S), V(p, O)\}$ and portfolio choices $\{y(p), g_J(p)\}$ are determined jointly. The state space consists of manager reputation in log-likelihood space ($p \in \mathbb{R}_+$).

I first discretize the state space as follows. I construct a grid with limits $\{0, \overline{p}\}$ and N grid points. Let $z = (\overline{p} + 1)^{1/N}$; I populate the grid by setting $p(j) = z^j - 1$. Using this grid, I discretize Equation (10) using central differences as described in Candler (1998).

First I hold the incentive distortion constant at zero, $y(p) = g_J(p) = 0$, and iterate to find the solution to the transparent portfolio problem $\{V(p, S), V(p, O)\}^{tr}$. The transparent liquidation threshold is a lower bound to the equilibrium liquidation threshold, because the expected return of the skilled expert is the highest. Starting from the transparent liquidation policy and value function, I iterate on equation (10), and each time I solve for the pair of choices $\{y, g_J\}$ at each step.

More specifically, the iteration procedure can be divided into the following steps.

1. Given $[y(p)]^{i-1}$ and $[g_J(p)]^{i-1}$, solve for $[X(P,\theta)]^i$, $[\sigma_a(p)]^i$, $[f(p)]^i$, $[\underline{P}]^i$, using Lemma 1 and Propositions 5, 6, and 7.
2. Solve for $[V(p, S), V(p, O)]^i$ using the discretization of Equation (10) and the boundary condition $V(0, \theta) = 0$.
3. Given $[V(p, S), V(p, O)]^i$ solve for $[y(p)]^i$ and $[g_J(p)]^i$ using Proposition 5 and Definition 1 (see discussion below).

4. If $|\,(V(p,S))^i-(V(p,S))^{i-1}\,|<\epsilon$ and $|\,(V(p,O))^i-(V(p,O))^{i-1}\,|<\epsilon$ are not satisfied, repeat steps 1 through 4.

The procedure converges extremely rapidly and takes no more than five iterations for a typical solution. The typical solution time is 30 seconds. Code is available upon request.

C.1. Fixed point

Step 3 consists of finding the fixed point of equation (14) that is consistent with the fastest rate of learning. For each reputation p_j grid value, I construct a grid with limits $\{\underline{y},1\}$ and N_π grid points, where $\underline{y}$ is a low negative number. I populate the grid using equal spacing across grid points. This produces $N\times N_y$ $y_{j,l}$ values. I look for all values of $y_{j,l}$, where the expression

$$y_{j,l}-\mathcal{D}_S(p_j,y_{j,l})$$

switches sign. The term $\mathcal{D}_S$ is defined in Proposition 3 and 5.1. This step identifies all incentive functions consistent with equilibrium. Specifically, for each grid point p_j, there is a set Ω_j that includes all incentives $y_{j,l}$ for which $y_{j,l}=\mathcal{D}_S(p_j,y_{j,l})$. If for a given reputation p_j the set Ω_j is a singleton, this incentive $y_{j,l}\in\Omega_j$ defines the equilibrium for reputation p_j, that is, $y_j=y$, where $y\in\Omega_j$. If there is more than one $y_{j,l}\in\Omega_j$, then I compute

$$y_j=\arg\max_{y\in\Omega_j}\left(\alpha_s\mathcal{X}(y)\mu_S-\lambda g_J(p_j,y)\right)^2+\lambda\left(\frac{g_J(p_j,y)}{\sqrt{\phi}}\right)^2 \tag{C.59}$$

according to Definition 1 and select the one with the largest learning rate.

Appendix D. Assumptions

A key modeling assumption is that intermediaries face a investment opportunity set that is rich enough to allow some reach-for-yield, but not too rich to allow unbounded gambling. This fits well with most of the investment management industry, where managers are constrained by investment mandates to a varying degree. Managers cannot trade arbitrary payoffs without raising red flags, but can easily distort their portfolio toward assets with high tail-risk exposure that are otherwise very similar. Coval et al. (2009) provide a useful example of such differences within the fixed income market. The assumption that tail risk cannot be perfectly measured even *ex-post* is realistic and also theoretically important. It avoids the situation where small performance differences during tail events are perfectly informative about the intermediary portfolio. This would be unrealistic and an artificial by-product of the continuous time environment. It would also introduce a lot of new equilibria. Intuitively, as the signal-to-noise ratio grows to infinity, all managers have strong incentives to conform exactly with investors expectations. In this sense, tail-risk volatility does something similar to the effect of private information in the global games literature.

Another critical assumption is that intermediaries can only use performance to signal their type. The assumption that there are no other signaling mechanisms is likely to be counter-factual in most environments. However, all that is required for the model's fundamental mechanism to work is some asymmetric information about intermediary skill between investors and the intermediary. If this is the case, investors will use performance to learn about intermediary type, and the link between capital immobility and reach for yield will exist. This is obviously realistic and consistent with the positive relation between past performance and flows present in the data for most intermediaries.

The assumption that there are only two intermediary types is stylized, but necessary for tractability in an environment where asymmetric information is persistent. Previous studies on the interaction of learning and investment decisions in money management (for example, Makarov and Platin, 2015) work in an environment where information about manager quality is symmetric across investors and managers. The assumption that the manager knows more about his quality is key to generating the link between reaching for yield and slow-moving capital. It is essential that managers who are observationally identical (same reputation), behave differently as a function of their privately known skill. This produces variation in reaching-for-yield incentives and renders performance less informative.

The contractual environment fits most mutual fund industry contracts well, and is consistent with the leading work-horse model in the money management literature (Berk and Green, 2004). However, it is admittedly stylized and abstracts from many realistic features present in the contracts of more sophisticated financial intermediaries, such as hedge funds.

Before discussing how the introduction of performance contracts would change the model results, it is useful to contrast the incentive problem that arises here with the one that shows up in the moral-hazard based models (for example, He and Krishnamurthy, 2013). There, performance incentives are needed to induce work or avoid tunneling by the intermediary. Here, it is driven by reputation incentives that arise endogenously from the sensitivity of the intermediary's human capital to his own performance track record. Provision of full insurance to the intermediary would implement the first best, but it is not feasible if there is competition among investors. Intuitively, investors outside of the relationship would make outside offers to intermediaries that are performing well, unraveling the insurance scheme. The fundamental force is inalienability of human capital (Hart and Moore, 1994): the intermediary cannot commit to supply human capital in the future at below-market price.

The introduction of symmetric performance fees has no impact on the investment behavior of the opportunistic type, but pushes the skilled experts toward the Sharpe portfolio. Intuitively, no matter his portfolio choice the opportunistic expert generates zero expected excess return, so a linear performance schedule does not influence his choice. Thus, linear incentive fees have the effect of further *slowing down* the flow of capital into the skilled intermediary, but have the positive effect of improving the efficiency of the static capital allocation by the skilled type, and therefore lead to higher expected returns.

Nonlinear contracts could potentially increase the speed of capital flows if they induce incentives on the opportunist type to perform well during a tail event. If the contract is sufficiently nonlinear, it could induce gambling on the opportunistic type, counter-balancing the reputation incentives. In practice, the typical nonlinear contract used among financial intermediaries, the high-water-mark contract, is likely to have a mixed effect on incentives. When the fund is sufficiently far from the high-water-mark, the contract places more weight on large return realizations. This would nudge the intermediary in the right direction. But when the manager is close to the high-water-mark, the contract is linear on small return realizations, but places less weight on really bad return realizations. This pushes the manager choice in the wrong direction.

Appendix E. Parameter choice

The model has four key parameters – μ_S^+, κ_+, ϕ, and λ. The empirical plausibility of the parameter choice used in the text is discussed below.

The scarcity of capital is measured by the Sharpe ratio of the skilled intermediary before transaction costs and fees, μ_S^+. I calibrate $\mu_S^+ S = 1$, which is reasonable since this matches the Sharpe ratio of successful hedge fund managers. The average intermediary generates much less than this value. The model's quantitative implications for the speed of capital flows are stronger for lower values of μ_S^+, holding $\lambda\kappa^+$ constant, and are unchanged if κ^+ decreases proportionally to μ_S^+. The intuition for this result is that what matters for the reduction in the speed of capital flows is how much of the normal-times performance advantage of the skilled type (μ_S^+) can be matched by the opportunist taking on tail risk. This can be seen clearly in Equation (8).

The second key parameter is κ^+, which is the maximum tail exposure a portfolio of unit variance can achieve. I set κ^+ to 2, which implies that in a portfolio with 10% standard deviation, the intermediary can find (hidden) opportunities to have a tail loss of 20%. Given the recent losses experienced in some fixed-income markets, this tail exposure seems plausible. For example, many of the economic catastrophic bonds of Coval et al. (2009) dropped to almost zero in the aftermath of the crisis. This parameter should strongly vary by asset class and investment mandate. What is important is that $\lambda\kappa^+$ is of the same order of magnitude as μ_S^+. For example, in this calibration $\lambda\kappa^+ = \frac{\mu_S^+}{2}$. If κ^+ is much lower than this, then the effects on the speed of capital flows are small. Intuitively, capital immobility is a direct result of the flexibility that the opportunistic type has in taking on tail risk.

Tail-risk volatility determines how informative tail-event performance is. My baseline calibration uses $\sqrt{\phi} = 3$, which is in line with recent experience. For example, some measures of realized stock market volatility in the fall of 2008 were 80%. This compares to an average market volatility of 15%. In fixed-income markets, this difference is likely to be even larger. As ϕ grows, the economy converges to the case of Proposition 9. As ϕ shrinks to zero, tail-event performance becomes completely revealing about the intermediary portfolio. Such a parametrization greatly increases the scope for multiple equilibria. More on this topic is discussed in Appendix D.

Tail event frequency λ determines the degree to which reaching-for-yield can boost performance during normal-times, but also determines how quickly the tail event arrives. The first effect increases the temptation of the intermediary to reach for yield, while the second effect acts in the opposite direction, as a disciplining force. I calibrate λ to 0.25, consistent with a tail event every four years. This number is consistent with recent experience. For example, in the past 20 years the world economy experienced the "Tequilla crisis" (1994), the "Asian crisis" (1997-8), the "Tech bubble" bust (2000), the more recent financial crisis (2007-2009), and the European sovereign debt crisis (2011-2012).

Other less important parameters include ρ, ψ_0, ψ_1, and $\mu_S^\top \Sigma\kappa$. I set $\rho = 0.05$. I set ψ_0 to target a liquidation threshold of 0.5, but this choice is arbitrary. The model works identically for any positive choice of ψ_0, only changing the liquidation threshold and the long-run size of the fund. It is important that $\psi_0 > 0$ so that managers face liquidation risk. The decreasing-returns-to-scale parameter ψ_1 is set so that the amount of capital allocated to the intermediary peaks at 1. This is also an arbitrary choice.

I assume $\mu_S^\top \Sigma\kappa = 0$ and $\mu_S \propto 1_n$ throughout most of the paper. This implies that the Sharpe and Tail-risk portfolios are exactly orthogonal and expected returns are constant across technologies. Economically this means that the Sharpe portfolio has zero tail exposure and the price of tail risk is zero. These assumptions control how much more yield the expert can find in her opportunity set relative to the Sharpe portfolio. For the choice $\mu_S^\top \Sigma\kappa = 0$, this difference is given by κ^+. For the more general case, this would be given by

$$\kappa^{+} - \frac{\mu_S^{\top} \Sigma^{-1} \kappa}{\sqrt{\mu_S^{\top} \Sigma^{-1} \mu_S}}. \tag{E.60}$$

Both an increase in the average tail risk across assets and an increase in the price of tail risk have the effect of reducing the spread in tail risk across portfolios. As discussed before, the results of the speed of capital flows rest on the assumption that the difference in tail exposure between the Sharpe portfolio and the maximum tail exposure feasible are economically meaningful.

References

Acharya, Viral V., Pagano, Marco, Volpin, Paolo, 2013. Seeking alpha: excess risk taking and competition for managerial talent. NBER Macroecon. Annu.

Basak, S., Pavlova, A., Shapiro, A., 2007. Optimal asset allocation and risk shifting in money management. Rev. Financ. Stud.

Basak, Suleyman, Pavlova, Anna, 2013. Asset prices and institutional investors. Am. Econ. Rev. 103 (5), 1728–1758.

Basak, Suleyman, Makarov, Dmitry, 2014. Strategic asset allocation in money management. J. Finance 69 (1), 179–217.

Berk, J.B., Green, R.C., 2004. Mutual fund flows and performance in rational markets. J. Polit. Econ. 112 (6), 1269–1295.

Berk, J.B., Stanton, R., 2007. Managerial ability, compensation, and the closed-end fund discount. J. Finance 62 (2), 529–556.

Berk, Jonathan B., Van Binsbergen, Jules H., 2015. Measuring skill in the mutual fund industry. J. Financ. Econ. 118 (1), 1–20.

Bessembinder, Hendrik, Kahle, Kathleen M., Maxwell, William F., Xu, Danielle, 2009. Measuring abnormal bond performance. Rev. Financ. Stud. 22 (10), 4219–4258.

Brennan, Michael J., Li, Feifei, 1993. Agency and Asset Pricing.

Candler, Graham V., 1998. Finite-Difference Methods for Continuous-Time Dynamic Programming. In: QM&RBC Codes, Quantitative Macroeconomics & Real Business Cycles.

Chapman, D.A., Evans, R.B., Xu, Z., 2009. Career Concerns and the Active Fund Manager's Problem. Working paper.

Chevalier, J., Ellison, G., 1997. Risk taking by mutual funds as a response to incentives. J. Polit. Econ. 105 (6), 1167–1200.

Choi, Jaewon, Kronlund, Mathias, 2014. Reaching for Yield or Playing it Safe? Risk Taking by Bond Mutual Funds. Working paper.

Coval, Joshua D., Jurek, Jakub W., Stafford, Erik, 2009. Economic catastrophe bonds. Am. Econ. Rev. 99 (3), 628–666.

Cuoco, Domenico, Kaniel, Ron, 2011. Equilibrium prices in the presence of delegated portfolio management. J. Financ. Econ. 101 (2), 264–296.

Dangl, T., Wu, Y., Zechner, J., 2008. Market discipline and internal governance in the mutual fund industry. Rev. Financ. Stud. 21 (5), 2307.

Dasgupta, Amil, Prat, Andrea, 2008. Information aggregation in financial markets with career concerns. J. Econ. Theory 143 (1), 83–113.

Dasgupta, Amil, Prat, Andrea, Verardo, Michela, 2011. The price impact of institutional herding. Rev. Financ. Stud. 24 (3), 892–925.

Di Maggio, Marco, 2014. Fake Alphas, Tail Risk and Reputation Traps. Working Paper.

Duffie, Darrell, 2010. Presidential Address: asset price dynamics with slow-moving capital. J. Finance 65 (4), 1237–1267.

Goetzmann, W.N., Ingersoll Jr., J.E., Ross, S.A., 2003. High-water marks and hedge fund management contracts. J. Finance 58 (4), 1685–1717.

Goldstein, Itay, Jiang, Hao, Ng, David T., 2015. Investor Flows and Fragility in Corporate Bond Funds. Available at SSRN 2596948.

Guerrieri, Veronica, Kondor, Peter, 2012. Fund managers, career concerns, and asset price volatility. Am. Econ. Rev. 102 (5), 1986–2017.

Hart, Oliver, Moore, John, 1994. A theory of debt based on the inalienability of human capital. Q. J. Econ. 109 (4), 841–879.

He, Zhiguo, Krishnamurthy, Arvind, 2013. Intermediary asset pricing. Am. Econ. Rev. 103 (2), 732–770.

Holmstrom, B., 1999. Managerial incentive problems: a dynamic perspective. Rev. Econ. Stud. 66 (1), 169–182.

Kacperczyk, Marcin, Schnabl, Philipp, 2012. How Safe are Money Market Funds?.

Kaniel, Ron, Kondor, Péter, 2013. The delegated Lucas tree. Rev. Financ. Stud. 26 (4), 929–984.

Kaplan, Steven N., Schoar, Antoinette, 2005. Private equity performance: returns, persistence, and capital flows. J. Finance 60 (4), 1791–1823.

Kelly, Bryan, Jiang, Hao, 2014. Tail risk and asset prices. Rev. Financ. Stud. 27 (10), 2841–2871.

Kozlowski, Julian, Veldkamp, Laura, Venkateswaran, Venky, 2016. The Tail that Wags the Economy: Belief-Driven Business Cycles and Persistent Stagnation.

Makarov, Igor, Platin, Guillaume, 2015. Rewarding trading skills without inducing gambling. J. Finance 70 (3), 925–962.

Ma, Linlin, Tang, Yuehua, Gómez, Juan-pedro, 2019. Portfolio manager compensation in the US mutual fund industry. J. Finance 74 (2), 587–638.

Malliaris, Steven G., Yan, Hongjun, 2015. Reputation Concerns and Slow-Moving Capital. Working paper.

Pastor, Lubos, Stambaugh, Robert F., 2010. On the size of the active management industry. NBER Macroecon. Annu.

Pástor, L'uboš, Stambaugh, Robert F., Taylor, Lucian A., 2015. Scale and skill in active management. J. Financ. Econ. 116 (1), 23–45.

Pedersen, Lasse Heje, Mitchell, Mark, Pulvino, Todd, 2007. Slow moving capital. Am. Econ. Rev. 97 (2), 215–220.

Rajan, Raghuram, 2008. Bankers pay is deeply flawed. Financial Times 8.

Rajan, Raghuram, 2012. Fault Lines. HarperCollins Publishers.

Rajan, Raghuram G., 2005. Has financial development made the world riskier? NBER Macroecon. Annu.

Shleifer, Andrei, Vishny, Robert W., 1997. The limits of arbitrage. J. Finance 52 (1), 35–55.

Stein, J., 2013. Overheating in Credit Markets: Origins, Measurement, and Policy Responses.

Vayanos, Dimitri, 2004. Flight to quality, flight to liquidity, and the pricing of risk. NBER Macroecon. Annu.

Vayanos, Dimitri, Woolley, Paul, 2013. An institutional theory of momentum and reversal. Rev. Financ. Stud. 26 (5), 1087–1145.