A Note on Additional Materials for "Common Fund Flows: Flow Hedging and Factor Pricing"

Winston Wei Dou   Leonid Kogan   Wei Wu*

December 30, 2023

Abstract

This note contains an extended version of the simple model in the paper titled "Common Fund Flows: Flow Hedging and Factor Pricing" (Dou, Kogan and Wu, 2023*a*). Section 1 describes the extended model. It is a discrete-time, infinite-horizon, overlapping-generations (OLG), general-equilibrium framework with multiple risky assets, one risk-free asset, and a single perishable consumption good. We postulate a simple specification of compensation contracts for fund managers, which is strongly supported in the data. In particular, we specify the compensation structure of fund managers based on the estimations by Cen et al. (2023) using US data, and Ibert et al. (2018) using Swedish data. Section 2 provides proofs for the theoretical results pertaining to the extended model.

Keywords: Agency conflicts, Mutual fund flows, Intermediary asset pricing, Heterogeneous agents, Uncertainty, Demand shocks. (**JEL:** G11, G12, G23)

*Dou is at University of Pennsylvania (Wharton) and NBER. Kogan is at MIT (Sloan) and NBER. Wu is at Texas A&M University (Mays). Emails: wdou@wharton.upenn.edu, lkogan@mit.edu, and wwu@mays.tamu.edu. We thank Ziyuan Lan for excellent research assistance. All errors are our own.

Contents

1 Extended Model

1.1 Assets

There are n risky assets in the economy, indexed by $i = 1, \cdots, n$. Their dividends are stacked in a n-dimensional vector $D_t = [D_{1,t}, \cdots, D_{n,t}]^T$, and the log dividends are $d_t = \ln(D_t)$. The data-generating process of the log dividend growth rates is

$$\Delta d_{t+1} = \mu + \sqrt{h_t}\left(Bu_{t+1} + \varepsilon_{t+1}\right), \tag{1}$$

where $u_t = [u_{1,t}, \cdots, u_{k,t}]^T$ are k primitive factors distributed as i.i.d. $N(0, I_k)$, and $\varepsilon_t = [\varepsilon_{1,t}, \cdots, \varepsilon_{n,t}]^T$ are residuals distributed as i.i.d. $N(0, I_n)$. The $n \times k$ matrix B captures the loading coefficients of the n log dividend growth rates Δd_{t+1} on the k factors u_{t+1}.

By postulating distributional structure (1) for log dividend growth, we assume that the covariance matrix of assets' cash flows is mainly captured by that of a few dominant factors, similar to many other multi-asset portfolio choice and asset pricing models (e.g., Kozak, Nagel and Santosh, 2018; Koijen and Yogo, 2019). This assumption is consistent with the empirical evidence documented by Ball, Sadka and Sadka (2009), who show that there is a strong factor structure in firms' fundamentals.

We assume that the number of assets in this economy, n, is large, and various cross-sectional averages of idiosyncratic shocks, e.g., $(1/n)\sum_{i=1}^{n}\varepsilon_i$, are approximately equal to 0, which is essentially the assumption of the Arbitrage Pricing Theory (e.g., Ross, 1976). In particular, the number of assets is much larger than the number of primitive factors, i.e., $1 \leq k \ll n$.

The time-varying uncertainty is characterized by univariate state variable h_t, which is driven by k aggregate shocks u_t as follows:[1]

$$h_{t+1} = \bar{h} + \rho(h_t - \bar{h}) + \sqrt{h_t}\sigma u_{t+1}, \quad \text{with } \rho \in (0,1) \text{ and } \sigma \in \mathbb{R}^{1\times k}. \tag{2}$$

[1] We impose a zero lower bound on h_t similar to Bansal and Yaron (2004), Chen, Dou and Kogan (2021), and Cheng, Dou and Liao (2022).

Without loss of generality, we assume that the $1 \times k$ vector $\sigma = [\sigma_1, \cdots, \sigma_k]$ has positive elements, i.e., $\sigma_j > 0$ for $j = 1, \cdots, k$.

Stock i is a claim to dividend stream $D_{i,t}$ for $i = 1, \cdots, n$, and is in unit net supply. Similar to Kozak, Nagel and Santosh (2018), we assume that the supply of the risk-free bond is perfectly elastic, with a constant risk-free rate of $R_f > 1$.[2] Let $r_f = \ln\left(R_f\right)$ denote the log risk-free interest rate. The return of risky asset i is given by $R_{i,t+1} \equiv (P_{i,t+1} + D_{i,t+1})/P_{i,t}$ where $P_{i,t}$ is the price of risky asset i at time t for $i = 1, \cdots, n$. The vector that stacks the risky asset returns is denoted by $R_{t+1} = [R_{1,t+1}, \cdots, R_{n,t+1}]^T$.

Log-Linear Approximation. We use a log-linear approximation to characterize the equilibrium relation among consumption, portfolio holdings, and asset prices analytically. The log return vector, $r_{t+1} \equiv \ln(R_{t+1})$, can be expressed as

$$r_{t+1} \approx L z_{t+1} - z_t + \Delta d_{t+1} + \ell, \tag{3}$$

where $z_t = \ln\left(P_t/D_t\right)$ is the $n \times 1$ vector of log price-dividend ratios with elements $z_{i,t} = \ln(P_{i,t}/D_{i,t})$. The matrix L in (3) is a $n \times n$ diagonal matrix with the ith diagonal element equal to $L_i = e^{\bar{z}_i}/(1 + e^{\bar{z}_i}) \in (0, 1)$, where $\bar{z}_i$ is the long-run average of the log price-dividend ratio for asset i. The vector ℓ in (3) is a $n \times 1$ vector with the ith element equal to $\ell_i = -\ln\left(L_i\right) + (1 - L_i)\ln\left(1/L_i - 1\right)$.

We conjecture that the log price-dividend ratio is an affine function of the aggregate state variable h_t:

$$z_t \approx \zeta + \zeta_h (h_t - \bar{h}), \tag{4}$$

where $\zeta, \zeta_h \in \mathbb{R}^{n \times 1}$ are constant vectors to be determined in equilibrium.

Based on the representation of log returns in (3) and the equilibrium log price-dividend ratio

[2] We fix the risk-free rate in the model for tractability. This assumption is not unreasonable for the US market, where US Treasuries are largely held and traded by foreign investors, and the risk-free rate is not determined entirely by domestic demand (e.g., Gourinchas and Rey, 2007; Caballero, Farhi and Gourinchas, 2008; Dou and Verdelhan, 2017).

in (4), equilibrium log returns r_{t+1} can thus be characterized as follows. The proof is in Section 2.1.

Proposition 1.1 (Excess log returns of risky assets)**.** *The equilibrium excess log returns of risky assets are*

$$r_{t+1} - r_f \mathbf{1} \approx \mu_t + \sqrt{h_t}\left(K u_{t+1} + \varepsilon_{t+1}\right), \tag{5}$$

where $\mathbf{1} \in \mathbb{R}^{n\times 1}$ *is a vector of ones,* $\mu_t \in \mathbb{R}^{n\times 1}$ *is the conditional expected excess log return given the information set up to time* t*, and* $K \in \mathbb{R}^{n\times k}$ *captures stock returns' systematic risk exposure:*

$$\mu_t = (\rho L - I_n)\zeta_h h_t \;\; and \;\; K = L\zeta_h \sigma + B, \tag{6}$$

where B *is defined in (1),* ρ *and* σ *are defined in (2),* L *is defined in (3), and* ζ_h *is defined in (4). The variance-covariance matrix of the log returns is*

$$\Sigma_t = \Sigma h_t, \;\; with \;\; \Sigma = I_n + KK^T. \tag{7}$$

In principle, factor models can arise in the equilibrium whether expected returns reflect systematic risk or mispricing. The macro factors u_{t+1} can capture systematic risks for which investors require compensation, or they can capture common sources of mispricing, such as market-wide investor sentiment (e.g., Hirshleifer and Jiang, 2010; Stambaugh and Yuan, 2016; Kozak, Nagel and Santosh, 2018).[3] Particularly, if the kth column of the loading matrix B in (1) is zero and the kth element of σ in (2) is strictly positive, the macro factor u_k tends to be a non-fundamental one (e.g., a sentiment factor or "mispricing factor").

Next, we approximate the portfolio's log return. Let $r_{t+1}(\phi) = \ln\left[R_{t+1}(\phi)\right]$ denote the log return of the portfolio with weights $\phi \in \mathbb{R}^{n\times 1}$. Then, we approximate the portfolio's log return as

$$r_{t+1}(\phi) \approx r_f + \phi^T(r_{t+1} - r_f\mathbf{1}) + \frac{1}{2}\phi^T\left(v_t - \Sigma_t\phi\right), \tag{8}$$

[3]Moreover, as emphasized, for example, by Long et al. (1990), there need not be a clear-cut distinction between mispricing and risk compensation as alternative justifications for multi-factor models of expected return. Specifically, Long et al. (1990) show that fluctuations in market-wide sentiment of noise traders give rise to a source of systematic risk for which rational traders require compensation.

where $v_t \equiv \text{diag}(\Sigma_t)$ is the vector that contains the diagonal elements of Σ_t.

1.2 Funds

To focus on the common component of fund flow shocks, we assume that the funds are homogenous.[4] The funds are typically active mutual funds and pension management, while fund clients are typically individual investors and pension sponsors. Funds can trade all assets freely, and they charge an advisory fee from fund clients. The advisory fee is a constant $f > 0$ fraction of AUM.[5]

Similar to the framework of Berk and Green (2004), we assume the active funds have skillful managers and information advantages to add value by generating ex-ante expected excess return relative to passive investment strategies. As argued by the literature (e.g., Vayanos and Woolley, 2013; Berk and van Binsbergen, 2015, 2016*a*; Pedersen, 2018; Leippold and Rueegg, 2020), there are some meaningful ways for active funds to outperform (i.e., add value) as a group.[6] Specifically, the value extracted by a mutual fund from the capital markets can be conceptualized as a wealth transfer from passive to active funds, which can occur in at least three distinct ways.

First, active fund managers can operate as informed arbitrageurs, capitalizing on new information to earn returns at the expense of uninformed investors, particularly index funds. This dynamic is grounded in the theoretical framework established by Grossman and Stiglitz (1980) and García and Vanden (2009).

Second, because index funds are compelled to closely track their benchmark indices, they inherently generate demand for immediacy in trading, often incurring associated costs. Active fund managers, unbound by such strict index-tracking mandates, have the opportunity to sidestep these costs of immediacy. Moreover, they have the potential to act as liquidity providers, positioning themselves to earn additional returns. Third, benchmark indices do not contain all available assets in the markets such as frontier markets, emerging markets, and private markets. This provides

[4]Heterogenous funds have been considered in studies on cross-fund flows (e.g., Berk and Green, 2004; Barber, Huang and Odean, 2016; Berk and van Binsbergen, 2015; Roussanov, Ruan and Wei, 2021).

[5]Different from Berk and Green (2004) and Kaniel and Kondor (2013), we assume exogenous constant expense ratio f for simplicity. The expense ratio can be endogenized similar to Kaniel and Kondor (2013).

[6]The authors show that the argument claiming it to be impossible for the average active fund manager to add value in a fully rational equilibrium (Sharpe, 1991) relies upon extremely strong assumptions.

ample scope for active fund managers to diverge from benchmark indices and explore profitable investment opportunities (e.g., Vayanos and Woolley, 2013).

Consider an active fund with Q_t assets under management (AUM). The value added by this fund is represented in reduced form by αQ_t, which is presumed to be independent of the fund's specific portfolio composition. The expected excess return, denoted by α, reflects the fund's gross alpha prior to the deduction of expenses and fees. Active funds incur various costs, which we assume to be increasing and convex in the AUM of the fund, as in Berk and Green (2004). Specifically, an active fund of size Q_t incurs a total cost of $\Psi(q_t)W_t$, where W_t is the total wealth of all agents, $q_t = Q_t/W_t$, and

$$\Psi(q) \equiv \theta^{-1} q^{1+\xi}, \text{ with } \xi > 0 \text{ and } \theta > 0. \tag{9}$$

Our specification essentially implies decreasing returns to scale for active funds.

The literature has advanced two hypotheses regarding the nature of the convex operating cost. The first one is fund-level decreasing returns to scale: as the size of an active fund increases, the fund's ability to outperform its benchmark declines (e.g., Perold and Salomon, 1991; Berk and Green, 2004). The second hypothesis is industry-level decreasing returns to scale: as the size of the active mutual fund industry increases, the ability of any given fund to outperform declines (Pástor and Stambaugh, 2012; Pástor, Stambaugh and Taylor, 2015). Both hypotheses are motivated by the price impact of trading and they are not mutually exclusive. At the fund level, a larger fund's trades have a larger impact on asset prices, eroding the fund's performance. At the industry level, as more money chases opportunities to outperform, prices move, making such opportunities more elusive. Consistent with such price impact of trading, there is mounting evidence showing that trading by mutual funds can exert meaningful price pressure in equity markets. Edelen and Warner (2001) and Ben-Rephael, Kandel and Wohl (2011) find that aggregate flow into equity mutual funds has an impact on aggregate market returns. Coval and Stafford (2007), Edmans, Goldstein and Jiang (2012), Khan, Kogan and Serafeim (2012), and Lou (2012) also find significant firm-level price impact associated with mutual fund trading. Edelen, Evans and Kadlec (2007) argue that trading costs are a major source of diseconomies of scale for mutual funds.

The expected excess total payout by the active funds to their clients is

$$TP_t = \overbrace{\overline{\alpha} Q_t - \Psi(q_t) W_t}^{\text{net gain of funds}} - f Q_t, \tag{10}$$

where $\overline{\alpha} Q_t$ is the value added by the active funds, $\Psi(q_t) W_t$ is the cost incurred by the active funds to create the gross alpha, and $f Q_t$ is the management fee charged by the active fund in period t.

We define the net alpha as $\alpha_t \equiv \frac{TP_t}{Q_t}$, which is the expected return received by the fund clients in period t in excess of the benchmark return:

$$\alpha_t = \overline{\alpha} - \psi(q_t) - f, \tag{11}$$

where $\psi(q_t) \equiv \Psi(q_t)/q_t = \theta^{-1} q_t^{\xi}$. We assume that $\xi = 1$ for the rest of this paper, and thus, the relation (11) can be rewritten as a linear relation between the amount of asset management service supplied by funds and the net alpha:

$$q_t = \theta(\overline{\alpha} - \alpha_t) - \theta f. \tag{12}$$

1.3 Agents

Different Types of Agents. The economy is populated by three different types of agents: direct investors, fund clients, and active fund managers. All investors can invest in and trade the risk-free asset. Direct investors, labeled by d, have to trade risky assets directly on their own accounts or hold passive investments such as benchmark indices; they are mainly index funds, passive exchange-traded funds (ETFs), and individual retail investors. Fund clients, labeled by c, have to delegate their risky-asset investments to professional active fund managers.[7] Fund clients can be retail individual investors or institutional investors such as pension sponsors or university

[7] This is a simplification. In the online appendix, we present an extended model in which fund clients can choose to trade risky assets directly.

endowments (e.g., Gerakos, Linnainmaa and Morse, 2021). Active fund managers, labeled by m, operate the funds, consume their fund revenues, and can invest in the risk-free asset on their own accounts to smooth consumption over time.

All agents live for two periods, forming overlapping generations (OLGs). Cohort-t agents are born in period t and die in period $t+1$. All agents have the same Epstein-Zin-Weil preference with a unitary elasticity of intertemporal substitution (EIS), a relative risk aversion (RRA) coefficient equal to γ, and subjective discount rate equal to β. Each agent in cohort t cares about her consumption in period t (when she is young) and the bequest to her descendants in period $t+1$ (when she is old). The utility function of agents of cohort t and type i is

$$U_{i,t} = (1-\beta)\ln\left(C_{i,t}\right) + \beta(1-\gamma)^{-1}\ln\mathbb{E}_t\left[\tilde{W}_{i,t+1}^{1-\gamma}\right], \quad \text{for } i \in \{d, c, m\}, \tag{13}$$

where $C_{i,t}$ and $\tilde{W}_{i,t+1}$ are cohort t's consumption and effective wealth in periods t and $t+1$, respectively.

A unit measure of newly born investors arrives at the beginning of each period. Investors are randomly assigned as fund clients with probability λ or as direct investors with probability $1-\lambda$. As a result, the newly-born direct investors are endowed with $(1-\lambda)W_t$ as their total initial wealth, while the newly-born fund clients are endowed with λW_t in total, where W_t is the total wealth of cohort t in period t. There is a unit measure of newly-born active fund managers with zero endowment.

We adopt an OLG framework to avoid tracking wealth shares as endogenous state variables when characterizing the equilibrium.[8] Moreover, we assume that agents in our model do not internalize their descendants' utility beyond the wealth term in Equation (13) to ensure that agents in our model are myopic.[9]

[8]Kaniel and Kondor (2013) showed how the constant wealth share of fund clients may arise endogenously as an equilibrium outcome. We can extend the model to endogenize the industry size of active equity funds, but we emphasize that it is not the focus of this paper to rationalize why a sizable industry of active equity funds would endogenously emerge as an equilibrium outcome. Rather, this paper explores how agency conflicts between active equity funds and their clients affect equity prices, given that active equity funds manage a large fraction of equity market investments.

[9]Seminal works (e.g., Barro, 1974; Abel, 1987) showed that OLG models with operative bequests are formally

Direct Investors. The direct investor's wealth is $W_{d,t} = (1-\lambda)W_t$. Direct investors solve a standard optimal portfolio problem. Denoting by $\phi_{d,t}$ the optimal portfolio weights of time t investable wealth $W_{d,t} - C_{d,t}$, we have

$$U_d(W_{d,t}) = \max_{\phi_{d,t}, C_{d,t}} (1-\beta)\ln(C_{d,t}) + \beta(1-\gamma)^{-1}\ln \mathbb{E}_t\left[W_{d,t+1}^{1-\gamma}\right], \tag{14}$$

subject to the dynamic budget constraint:

$$W_{d,t+1} = (W_{d,t} - C_{d,t} - \overline{\alpha}Q_t)\left[R_f + \phi_{d,t}^T(R_{t+1} - R_f)\right]. \tag{15}$$

Here, $\overline{\alpha}Q_t$ is the transfer of wealth from direct investors to active funds as discussed in Section 1.2.

Proposition 1.2 (Direct investors)**.** *The optimal consumption of direct investors is*

$$C_{d,t} = (1-\beta)(1-\lambda-\overline{\alpha}q_t)W_t, \tag{16}$$

and the optimal portfolio of direct investors is the standard myopic mean-variance efficient portfolio:

$$\phi_{d,t} = \frac{1}{\gamma}\Sigma_t^{-1}\left(\mu_t - r_f\mathbf{1} + \frac{1}{2}\nu_t\right), \tag{17}$$

where μ_t and Σ_t are defined in Proposition 1.1, and ν_t contains the diagonal elements of Σ_t.

See Section 2.2 for the proof in detail.

Fund Clients. Fund clients decide the amount of wealth to delegate to the funds, denoted by Q_t, and then the fund managers make allocation decisions for the delegated funds. Barber, Huang and Odean (2016) and Berk and van Binsbergen (2016*b*) find evidence that fund clients are not perfectly sophisticated in terms of incorporating the consideration of intertemporal hedging when they assess fund performance and make delegation decisions. To highlight this lack of sophistication,

equivalent to models with infinitely lived representative agents. Our assumption violates the conditions to ensure operative bequests. As a result, investors in our model are myopic.

we assume that fund clients behave myopically and do not hold rational expectations about funds' strategies. In particular, fund clients in our model do not properly anticipate that portfolios of fund managers depend on the delegation choice of the next generation of fund clients. Instead, we assume that fund clients care about the net alpha of the active managers relative to investing in the passive benchmark.[10] The fund clients are also free to become direct investors and manage their own portfolios.

We assume that the fund clients solve the following problem:

$$U_c(W_{c,t}) = \max_{C_{c,t},Q_t} (1-\beta)\ln(C_{c,t}) + \beta(1-\gamma)^{-1}\ln \mathbb{E}_t\left[(W_{c,t+1}+\omega Q_t)^{1-\gamma}\right], \tag{18}$$

subject to the budget constraint:

$$W_{c,t+1} = (W_{c,t} - C_{c,t})R_f + Q_t[R_{t+1}(\phi_{d,t}) + \alpha_t - R_f], \tag{19}$$

and the participation constraint:

$$U_c(W_{c,t}) \geq U_d(W_{c,t}). \tag{20}$$

The utility function in (18) contains the non-pecuniary benefit ωQ_t echoing the important insight that the net alpha in the eyes of a fund client depends on the client's specific utility of delegation (e.g., Ferson and Lin, 2014). And more specifically, the non-pecuniary benefit ωQ_t can be interpreted as the trust in active managers perceived by fund clients (Gennaioli, Shleifer and Vishny, 2015). The wealth evolution according to budget constraint (19) is intuitive. The fund client consumes $C_{c,t}$ out of wealth $W_{c,t}$, invests $W_{c,t} - C_{c,t} - Q_t$ to the risk-free bond, and delegates Q_t to the fund manager with perceived return $R_{t+1}(\phi_{d,t}) + \alpha_t$ and additional non-pecuniary benefit ωQ_t. The participation constraint (20) recognizes that fund clients are free to switch to direct investors, and it needs to hold to ensure that fund clients would decide to trust the active funds

[10]While we model the behavior of fund clients to be consistent with the main thrust of the recent literature on mutual fund flow, the precise behavioral assumptions we make are not essential for the key conclusions of our model about mutual fund hedging of common fund flow shocks, and the risk premium the flow-hedging demand generates. The essential element of the fund client's behavior is that they reduce their investment in equity mutual funds in high-uncertainty states when facing heightened economic uncertainty.

and delegate their investment management. When the term, ωQ_t, is sufficiently large, fund clients would choose to delegate their investment management even when the net alpha α_t is negative.

The following proposition characterizes the optimal consumption and delegation decision of fund clients.

Proposition 1.3 (Fund clients)**.** *If the perceived benefit from active management is sufficiently large relative to the cost of delegation, i.e.,* $\overline{\alpha} + \omega > \theta^{-1}\beta\lambda + f$*, fund clients choose to delegate their portfolios to the active funds. In this case, the optimal consumption of fund clients is*

$$C_{c,t} = (1-\beta)\lambda W_t, \tag{21}$$

and the total amount of asset management service demanded by fund clients satisfies

$$q_t = \beta\lambda\left(1 + \frac{\omega + \alpha_t}{\overline{\gamma} h_t}\right), \tag{22}$$

where ω *is the non-pecuniary benefit as in (18), and the term* $\overline{\gamma} h_t$ *captures the effective risk aversion with* $\overline{\gamma} \equiv \left[(\rho L - I_n)\zeta_h + \frac{1}{2}\nu\right]^T \Sigma^{-1}\left[(\rho L - I_n)\zeta_h + \frac{1}{2}\nu\right]$*, and* $\nu \equiv \mathrm{diag}(\Sigma)$*. Here* ρ*,* L*, and* ζ_h *are defined in (2), (3), and (4), respectively.*

In our theory, delegation to active funds is endogenously caused by (i) the net alpha of the active asset management α_t, (ii) the non-pecuniary benefit of the fund client, ω, and (iii) the degree to which the excess return incentivizes the investors to delegate their wealth to active asset management, captured by economic uncertainty h_t. The proof of Proposition 1.3 is in Section 2.3.

Active Fund Managers. The AUM of an active fund at the beginning of period t is Q_t and the revenue of the fund is advisory fee fQ_t. We assume that the fund manager of cohort t gets paid by fQ_{t+1} in period $t+1$, meaning that there is no agency conflict between the fund complex and the fund manager. A similar simplifying assumption has been commonly adopted in the literature.[11] What is crucial for our theoretical results is that active fund managers are concerned with their

[11] E.g., Brennan (1993), Gómez and Zapatero (2003), Basak, Pavlova and Shapiro (2007), Chapman, Evans and Xu (2010), Cuoco and Kaniel (2011), Kaniel and Kondor (2013), Basak and Pavlova (2013), and Koijen (2014).

fund's AUM, a fact that is robustly supported by data from the US (Cen et al., 2023) and Sweden (Ibert et al., 2018).[12]

Active fund managers in our model can save, and they don't have to consume fund revenues immediately period by period. But, importantly, we assume that active fund managers cannot invest in risky assets using their private wealth. This simplifying assumption has been widely adopted in the literature (e.g., Berk and Green, 2004; Cuoco and Kaniel, 2011; Kaniel and Kondor, 2013) for technical tractability, and enables us to avoid keeping track of active fund managers' private wealth, investment decisions, and associated constraints. Our theoretical results apply as long as the fund manager is unable to hedge against the flow risk fully by trading on a personal account for reasons such as liquidity constraints, leverage constraints, and frictions associated with short sales.

The active fund manager of cohort t chooses fund portfolio $\phi_{m,t}$ and consumption $C_{m,t}$ optimally to maximize the utility in Equation (13) subject to the budget constraint:

$$\tilde{W}_{m,t+1} = fQ_{t+1} - C_{m,t}R_f, \quad \text{with} \tag{23}$$

$$Q_{t+1} = \underbrace{Q_t\left[R_{t+1}(\phi_{m,t}) + \alpha_t\right]}_{\text{fund returns}} + \underbrace{Q_t flow_{t+1}}_{\text{fund flows}}, \tag{24}$$

where Q_t is the delegation characterized in Equation (22) given net alpha α_t and aggregate state h_t, and $Q_t flow_{t+1}$ is the net fund flow into the active fund.[13]

Equation (24) essentially gives the definition of the fund flow, denoted by $flow_{t+1}$:

$$flow_{t+1} \equiv \frac{Q_{t+1} - Q_t\left[R_{t+1}(\phi_{m,t}) + \alpha_t\right]}{Q_t}. \tag{25}$$

The dynamic budget constraint in Equation (24) is very intuitive. The total asset valuation

[12]Both papers document the fact that the compensation of individual fund managers is influenced by fund flow and return, primarily through their AUM (or revenue). Importantly, Cen et al. (2023) provides direct evidence from US data, indicating that the relationship between fund manager compensation and AUM (or revenue) is causal, that is, contractual in nature.

[13]We assume that active fund managers are myopic to highlight that our equilibrium results do not require any agents to engage in intertemporal hedging. This assumption is in fact consistent with active fund managers' short-term focus stemming from their career concerns (e.g., Prat, 2005; Hermalin and Weisbach, 2012).

at the beginning of period $t+1$ is $Q_t\left[R_{t+1}(\phi_{m,t})+\alpha_t\right]$ because active fund managers would consume management fees fQ_t and incur costs $\psi(q_t)Q_t$ to add value $\overline{\alpha}Q_t$ for active funds. The total AUM at the beginning of period $t+1$ is the sum of the fund return and fund flow: $Q_{t+1}=Q_t\left[R_{t+1}(\phi_{m,t})+\alpha_t+flow_{t+1}\right]$.

1.4 Equilibrium

Fund flow $flow_{t+1}$ and net alpha α_t after fees are endogenous, driven by aggregate shocks in a predictable way in equilibrium. Below, we describe how fund flows depend on fund managers' portfolio $\phi_{m,t}$ and aggregate shocks u_t.

Equilibrium Delegation and Endogenous Flows. Market clearing in the market for delegated funds is described by the two relations between the total amount of delegated capital and the net alpha – the first describing the alpha production technology of mutual funds, and the second describing the delegation decision of fund clients:

$$q_t=\theta(\overline{\alpha}-f)-\theta\alpha_t \quad \text{(funds' supply for asset management service)},$$
$$q_t=\beta\lambda\left(1+\frac{\omega+\alpha_t}{\overline{\gamma}h_t}\right) \quad \text{(clients' demand for asset management service)}.$$

Proposition 1.4 below summarizes the solution. The proof is in Section 2.4.

Proposition 1.4 (Equilibrium delegation and alpha). *The equilibrium amount of delegation q_t and the net alpha α_t are given by*

$$\alpha_t=-\omega+\frac{\theta(\overline{\alpha}+\omega-f)-\beta\lambda}{\theta+\beta\lambda/(\overline{\gamma}h_t)} \quad \textit{and} \quad q_t=\beta\lambda\left[1+\frac{\theta(\overline{\alpha}+\omega-f)-\beta\lambda}{\theta\overline{\gamma}h_t+\beta\lambda}\right], \tag{26}$$

where ω is the non-pecuniary benefit term in (18), $\overline{\gamma}$ is defined in (22), and gross alpha $\overline{\alpha}$, cost coefficient θ, and advisory fee f are defined in Section 1.2.

Corollary 1.1 (Countercyclical net alpha and pro-cyclical delegation). *When the benefits from active management are large relative to the cost of delegation, i.e., $\overline{\alpha}+\omega>\theta^{-1}\beta\lambda+f$, the equilibrium net alpha* *of funds is countercyclical and the equilibrium delegation is pro-cyclical. That is, α_t rises and q_t declines as uncertainty h_t increases:*

$$\frac{\partial \alpha_t}{\partial h_t} > 0 \quad \textit{and} \quad \frac{\partial q_t}{\partial h_t} < 0. \tag{27}$$

With the characterization of equilibrium delegation q_t, we are now ready to characterize the endogenous fund flows in equilibrium. We first conjecture the equilibrium aggregate fund flow

$$flow_{t+1} - \mathbb{E}_t\left[flow_{t+1}\right] \approx \sqrt{h_t} A u_{t+1}, \tag{28}$$

where $\mathbb{E}_t\left[flow_{t+1}\right] \in \mathbb{R}$ and $A \in \mathbb{R}^{1\times k}$ are to be determined in the equilibrium. According to (24), the process of fund flows can be approximated as shown in Proposition 1.5, whose proof is in Section 2.5.

Proposition 1.5 (Equilibrium aggregate fund flows)**.** *The exposure of common fund flows to the aggregate primitive shocks satisfies*

$$A = \frac{q'(\bar{h})}{q(\bar{h})}\sigma + \left[1 - \eta(\bar{h})\right] A K^T \left(I_n + KK^T\right)^{-1} K, \tag{29}$$

where function $q(h_t)$ is defined as in (26). And thus, the exposure of common fund flows to the aggregate primitive shocks is

$$A = \frac{q'(\bar{h})}{q(\bar{h})}\sigma \left\{ I_k - \left[1 - \eta(\bar{h})\right] K^T \left(I_n + KK^T\right)^{-1} K \right\}^{-1}, \tag{30}$$

where $\eta(\bar{h}) \equiv q(\bar{h}) / \left[(1-\lambda)\beta + (1-\overline{\alpha}) q(\bar{h})\right]$ and $\eta(h_t)$ captures the endogenous delegation intensity, which is derived in Theorem 2 below.

According to Corollary 1.1, each element of $\frac{q'(\bar{h})}{q(\bar{h})}\sigma$ is negative, which captures the negative relation between primitive shocks and changes in equilibrium delegation q_t, as well as the mechanical relation between fund flows and fund size, q_t. Because the $k \times k$ matrix $\left[1 - \eta(\bar{h})\right] K^T \left(I_n + KK^T\right)^{-1} K$ is positive definite, Proposition 1.5 shows that the flow-hedging portfolio held by the active fund managers has a dampening effect on the sensitivity of fund flows to primitive shocks in equilibrium (i.e., the magnitude of A decreases in $\eta(\bar{h})$). Meanwhile, the eigenvalues of $\left[1-\eta(\bar{h})\right] K^T \left(I_n + KK^T\right)^{-1} K$ are all between 0 and 1, and thus, exposure of fund flows to aggregate primitive shocks exists and is (approximately) equal to the quantity in (30). We emphasize the endogenous nature of fund flows, which is manifested by the fact that the endogenous steady-state delegation intensity is determined by the market clearing condition of competitive equilibrium illustrated in Theorem 2 below.

Theorem 1 shows that the optimal portfolio of the fund manager has two components — a myopic and a flow-hedging component. See Section 2.6 for proof.

Theorem 1 (Equilibrium fund portfolio)**.** *Fund managers hold a tilted portfolio to hedge against fluctuations in fund flows at the cost of a reduced Sharpe ratio:*

$$\phi_{m,t} = \phi_{d,t} - \phi_{\tau,t}, \tag{31}$$

where the optimal portfolio of the fund manager, $\phi_{m,t}$, is different from that of the direct investors, $\phi_{d,t}$ (i.e., the mean-variance efficiency portfolio), and the portfolio tilt of active fund $\phi_{\tau,t}$ is the hedging demand for the common fund flow:

$$\phi_{\tau,t} = \Sigma_t^{-1} \mathcal{B}_t. \tag{32}$$

Here, $\mathcal{B}_t \equiv Cov_t\left[r_{t+1}, flow_{t+1}\right]$ is the vector of fund flow betas, and in equilibrium, $\mathcal{B}_t = \mathcal{B} h_t$ with $\mathcal{B} \approx KA^T \in \mathbb{R}^{n\times 1}$. Subscript τ in $\phi_{\tau,t}$ stands for tilting.

The main theoretical result of this paper is that the portfolio tilt of the active fund relative to the benchmark is, on average, greater when the common fund flow beta is higher. We formalize this insight in Corollary 1.2, whose proof can be found in Section 2.7.

Corollary 1.2 (Portfolio tilt and common flow beta)**.** *The cross-sectional covariance between the two n-dimensional vectors $\mathcal{B}_t$ and $\phi_{\tau,t}$ is always positive:*

$$Cov\left[\mathcal{B}_t, \phi_{\tau,t}\right] > 0, \quad \textit{for each } t. \tag{33}$$

Competitive Equilibrium. Now we formally state the definition of the equilibrium. We focus on the symmetric competitive equilibrium with atomistic homogeneous fund managers, fund clients, and direct investors. Formally speaking, we are looking for a stationary symmetric competitive equilibrium defined as follows.

Definition 1.1 (Competitive equilibrium). *A competitive equilibrium is a price process, P_t, for the stocks, a risk-free rate, r_f, a fund's net alpha process, α_t, offered by the fund, consumption processes $C_{c,t}$ and $C_{d,t}$ of investors, and portfolio processes $\phi_{d,t}$, $\phi_{m,t}$, and q_t of investors such that*

(i) given the equilibrium prices, fund's excess return, and aggregate allocations,

(i.a) each direct investor's consumption $C_{d,t}$ and portfolio strategy $\phi_{d,t}$ are optimal in terms of maximizing the utility in (14) subject to (15);

(i.b) each fund client's consumption $C_{c,t}$ and delegation decision (portfolio strategy) q_t are optimal in terms of maximizing the utility in (18) subject to (19) and (20);

(i.c) each fund manager's portfolio strategy $\phi_{m,t}$ is optimal in terms of maximizing the utility in (13) subject to (23) and (24);

(ii) prices P_t, risk-free rate r_f, and fund's net alpha α_t clear goods, assets, and delegation markets:

(ii.a) goods market: $\sum_{i=1}^{n} D_{i,t} = C_{d,t} + C_{c,t} + fQ_t + \Psi(q_t)W_t$;

(ii.b) delegation market: $\psi^{-1}(\overline{\alpha} - \alpha_t - f) = q_t$;

(ii.c) assets market: $Q_t\phi_{m,t} + [W_{d,t} - C_{d,t} - \overline{\alpha}Q_t]\,\phi_{d,t} = [W_{d,t} - C_{d,t} + (1-\overline{\alpha})Q_t]\,\phi_t^{mkt}$.

The market clearing condition (ii.a) reflects that the total goods, $\sum_{i=1}^{n} D_{i,t}$ are either consumed by the agents (i.e., $C_{d,t} + C_{c,t} + fQ_t$) or used by the active fund managers to create gross alphas (i.e., $\Psi(q_t)W_t$). The market clearing condition (ii.b) is essentially the demand curve of delegation (12), and the supply curve of delegation (22) results from the optimization condition (i.b). The market clearing condition (ii.c) effectively characterizes the market portfolio in the economy, leading to the relation among the market portfolio, the myopic portfolio, and the active fund's portfolio, summarized in Theorem 2.

The great contribution of the CAPM theory is to connect systematic risk to return covariance with the market portfolio returns, which can be approximated in the data. Considering the deviation of active equity mutual funds' holdings $\phi_{m,t}$ from the market portfolio, ϕ_t^{mkt}, we can construct useful empirical tests for our fund flow hedging results. Specifically, the testable implication can be summarized in Theorem 2. See Section 2.8 for proof.

Theorem 2 (Portfolio tilt from the market portfolio and common flow beta)**.** *The fund managers hold a tilted portfolio to hedge against fluctuations in fund flows, relative to the market portfolio:*

$$\phi_{m,t} = \phi_t^{mkt} - (1-\eta_t)\phi_{\tau,t}, \tag{34}$$

where $\eta_t = \eta(h_t) \equiv q_t / \left[(1-\lambda)\beta + (1-\overline{\alpha})q_t\right] \in [0,1]$, and portfolio tilt of an active fund $(1-\eta_t)\phi_{\tau,t}$ is the additional hedging demand for the common fund flow relative to the market portfolio, with $\phi_{\tau,t}$ defined in (32). Thus, the cross-sectional covariance between the deviation of fund holdings from the market portfolio and the common flow beta is always negative:

$$Cov\left[\mathcal{B}_t, \phi_{m,t} - \phi_t^{mkt}\right] < 0, \quad \text{for each } t. \tag{35}$$

In equilibrium, common fund flows respond to aggregate economic shocks, and thus risk premia analogous to the hedging term in the ICAPM emerge even in a myopic environment, which is summarized in the following theorem, whose proof is in Section 2.9.

Theorem 3 (Conditional two-beta asset pricing model)**.** *For any portfolio $r_{t+1}(\phi) = \phi^T r_{t+1}$ with $\mathbf{1}^T\phi = 1$, the risk premium is explained by the covariance with the market return, denoted by $r_{t+1}(\phi_t^{mkt})$, and the covariance with the common fund flow, denoted by $flow_{t+1}$:*

$$\mathbb{E}_t\left[r_{t+1}(\phi)\right] - r_f + \frac{1}{2}\phi^T \nu_t \approx \underbrace{\gamma Cov_t\left[r_{t+1}(\phi),\ r_{t+1}(\phi_t^{mkt})\right]}_{\text{explained by market beta}} + \underbrace{\eta_t \gamma Cov_t\left[r_{t+1}(\phi),\ flow_{t+1}\right]}_{\text{explained by flow beta}},$$

where $\frac{1}{2}\phi^T\nu_t$ is the Jensen's term and η_t is defined in Theorem 2.

If $\text{Cov}_t\left[r_{t+1}(\phi),\ flow_{t+1}\right] < 0$, portfolio ϕ provides a natural hedging against fluctuations in the common fund flow.

Corollary 1.3 (CAPM holds when there is no delegation)**.** *When there is no delegation in the economy, i.e., $\lambda = 0$, Theorem 3 implies the conditional CAPM:*

$$\mathbb{E}_t\left[r_{t+1}(\phi)\right] - r_f + \frac{1}{2}\phi^T v_t \approx \gamma Cov_t\left[r_{t+1}(\phi),\ r_{t+1}(\phi_t^{mkt})\right]. \tag{36}$$

It further implies that the CAPM holds:

$$\mathbb{E}\left[r_{t+1}(\phi) - r_f + \frac{1}{2}\phi^T v_t\right] \approx \beta^{mkt}(\phi)\Lambda. \tag{37}$$

where $\beta^{mkt}(\phi) \equiv Cov\left[r_{t+1}(\phi),\ \hat{r}_{t+1}(\phi_t^{mkt})\right]/Var\left[\hat{r}_{t+1}(\phi_t^{mkt})\right]$ is the market beta with $\hat{r}_{t+1}(\phi_t^{mkt}) \equiv r_{t+1}(\phi_t^{mkt}) - \mathbb{E}_t\left[r_{t+1}(\phi_t^{mkt})\right]$, and $\Lambda \equiv \gamma\bar{h}\left[(\rho L - I_n)\zeta_h + \frac{1}{2}\nu\right]^T \Sigma^{-1}\left[(\rho L - I_n)\zeta_h + \frac{1}{2}\nu\right]$ is the market price of risk.

When there is no fund client in the economy (i.e., $\lambda = 0$), the equilibrium delegation is 0 (i.e., $q_t \equiv 0$) according to Proposition 1.4, leading to $\eta_t \equiv 0$. In this case, every investor consumes $C_t = (1-\beta)W_t$ and holds the mean-variance myopic portfolio $\phi_{d,t} = \frac{1}{\gamma}\Sigma_t^{-1}\left(\mu_t - r_f + \frac{1}{2}v_t\right)$. The proof of Corollary 1.3 is in Section 2.10.

Corollary 1.4 (Multifactor asset pricing)**.** *The primitive aggregate shocks are correlated with the common component of fund flows, so they are priced in the cross-section just as in the ICAPM framework:*

$$\mathbb{E}_t\left[r_{t+1}(\phi)\right] - r_f + \frac{1}{2}\phi^T v_t \approx \gamma Cov_t\left[r_{t+1}(\phi),\ r_{t+1}(\phi_t^{mkt})\right] + \sum_{j=1}^{k}\eta_t\gamma A_j\sqrt{h_t}Cov_t\left[r_{t+1}(\phi),\ u_{j,t+1}\right],$$

where $\frac{1}{2}\phi^T v_t$ is the Jensen's term, A_j is the j-th element of A, and η_t is defined in Theorem 2.

2 Proofs for the Extended Model

This section presents the proofs for the theoretical results pertaining to the extended model, which are parallel to those for the simple model presented in Dou, Kogan and Wu (2023*b*).

2.1 Proof for Proposition 1.1

According to (3), the log-linearization approximation leads to the following representation:

$$r_{t+1} \approx Lz_{t+1} - z_t + \Delta d_{t+1} + \ell. \tag{38}$$

Plugging (1) and (4) into the equation above, we can obtain

$$r_{t+1} \approx L(\zeta + \zeta_h(h_{t+1} - \bar{h})) - (\zeta + \zeta_h(h_t - \bar{h})) + \mu + \sqrt{h_t}Bu_{t+1} + \sqrt{h_t}\varepsilon_{t+1} + \ell. \tag{39}$$

Further, if we plug (2) into the relation above, we can obtain

$$r_{t+1} \approx L(\zeta + \zeta_h(\rho(h_t - \bar{h}) + \sqrt{h_t}\sigma u_{t+1})) - (\zeta + \zeta_h(h_t - \bar{h})) + \mu + \sqrt{h_t}Bu_{t+1} + \sqrt{h_t}\varepsilon_{t+1} + \ell. \tag{40}$$

Rearranging terms further leads to

$$r_{t+1} \approx \mathbb{E}_t\left[r_{t+1}\right] + \sqrt{h_t}Ku_{t+1} + \sqrt{h_t}\varepsilon_{t+1}, \tag{41}$$

where

$$\mathbb{E}_t\left[r_{t+1}\right] \approx \mu + \ell + (L - I_n)\zeta + (\rho L - I_n)\zeta_h(h_t - \bar{h}) \;\text{ and }\; K = L\zeta_h\sigma + B. \tag{42}$$

Moreover, if $h_t = 0$, all assets are risk-free during period t. Thus, the conditional expected returns in μ_t all equal risk-free rate r_f. Therefore, according to (6), the log risk-free rate must satisfy the following condition in the equilibrium to rule out arbitrage opportunities:

$$r_f\mathbf{1} \approx [\mu + \ell + (L - I_n)\zeta] - (\rho L - I_n)\zeta_h\bar{h}. \tag{43}$$

We now derive the expression for the conditional expected log excess return, which is approximately proportional to the stochastic variance h_t In fact, based on (6) and (43), it follows that

$$\mathbb{E}_t\left[r_{t+1}\right] - r_f\mathbf{1} \approx (\rho L - I_n)\zeta_h h_t. \tag{44}$$

2.2 Proof for Proposition 1.2

Plugging in the budget constraint, the optimization problem can be rewritten as

$$\max_{\phi_{d,t}, C_{d,t}} (1-\beta)\ln(C_{d,t}) + \beta\ln(W_{d,t} - C_{d,t} - \overline{\alpha}Q_t) + \beta(1-\gamma)^{-1}\ln\mathbb{E}_t\left\{\left[R_f + \phi_{d,t}^T(R_{t+1} - R_f)\right]^{1-\gamma}\right\}.$$

Thus, the unit EIS allows for the separation of optimal consumption and optimal portfolio problems. The optimal consumption is straightforward to derive:

$$C_{d,t} = (1-\beta)(W_{d,t} - \overline{\alpha}Q_t). \tag{45}$$

Following Campbell and Viceira (1999, 2001), we approximate the dynamic budget constraint $r_{t+1}(\phi_{d,t}) = \ln\left[R_{t+1}(\phi_{d,t})\right]$ as follows

$$r_{t+1}(\phi_{d,t}) \approx r_f + \phi_{d,t}^T(r_{t+1} - r_f\mathbf{1}) + \frac{1}{2}\phi_{d,t}^T\left(v_t - \Sigma_t\phi_{d,t}\right) \tag{46}$$

where $v_t \equiv \text{diag}(\Sigma_t)$ is the vector that contains the diagonal elements of Σ_t. The optimal portfolio problem becomes

$$\max_{\phi_{d,t}}(1-\gamma)^{-1}\ln\mathbb{E}_t\left\{e^{(1-\gamma)\left[r_f+\phi_{d,t}^T(r_{t+1}-r_f\mathbf{1})+\frac{1}{2}\phi_{d,t}^T\left(v_t-\Sigma_t\phi_{d,t}\right)\right]}\right\} \tag{47}$$

Using the moment generating function of multivariate normal variables, it follows that

$$\mathbb{E}_t\left\{e^{(1-\gamma)\left[r_f+\phi_{d,t}^T(r_{t+1}-r_f\mathbf{1})+\frac{1}{2}\phi_{d,t}^T\left(v_t-\Sigma_t\phi_{d,t}\right)\right]}\right\} = e^{(1-\gamma)\left[r_f+\phi_{d,t}^T(\mu_t-r_f\mathbf{1})+\frac{1}{2}\phi_{d,t}^T\left(v_t-\Sigma_t\phi_{d,t}\right)\right]+(1-\gamma)^2\frac{1}{2}\phi_{d,t}^T\Sigma_t\phi_{d,t}}$$

Thus, the optimal portfolio problem can be further rewritten as

$$\max_{\phi_{d,t}} \phi_{d,t}^T(\mu_t - r_f\mathbf{1} + \frac{1}{2}\nu_t) - \frac{\gamma}{2}\phi_{d,t}^T\Sigma_t\phi_{d,t}. \tag{48}$$

The first-order condition leads to

$$\phi_{d,t} = \frac{1}{\gamma}\Sigma_t^{-1}(\mu_t - r_f\mathbf{1} + \frac{1}{2}\nu_t). \tag{49}$$

2.3 Proof for Proposition 1.3

Denote $\phi_{c,t} \equiv \frac{Q_t}{W_{c,t} - C_{c,t}}$. Plugging in the budget constraint, the optimization problem can be rewritten as

$$\begin{aligned}\max_{\phi_{c,t}, C_{c,t}} & (1-\beta)\ln(C_{c,t}) + \beta\ln(W_{c,t} - C_{c,t}) & (50)\\ & + \beta(1-\gamma)^{-1}\ln\mathbb{E}_t\left\{\left[R_f + \phi_{c,t}\left(R_{t+1}(\phi_{d,t}) + \alpha_t + \omega - R_f\right)\right]^{1-\gamma}\right\},\end{aligned}$$

Thus, the unit EIS coefficient allows for the separation of optimal consumption and optimal portfolio problems. The optimal consumption is straightforward to derive:

$$\begin{aligned}C_{c,t} &= (1-\beta)W_{c,t} & (51)\\ &= (1-\beta)\lambda W_t. & (52)\end{aligned}$$

Following Campbell and Viceira (1999, 2001), we approximate the dynamic budget constraint $r_{\alpha,t+1}(\phi_{d,t}) = \ln\left[R_{t+1}(\phi_{d,t}) + \alpha_t + \omega\right]$ as follows

$$\begin{aligned}r_{\alpha,t+1}(\phi_{d,t}) &\approx \ln\left[R_{t+1}(\phi_{d,t}) + \alpha_t + \omega\right] & (53)\\ &\approx \alpha_t + \omega + r_f + \phi_{d,t}^T(r_{t+1} - r_f\mathbf{1}) + \frac{1}{2}\phi_{d,t}^T\left(v_t - \Sigma_t\phi_{d,t}\right), & (54)\end{aligned}$$

where $v_t \equiv \text{diag}(\Sigma_t)$ is the vector that contains the diagonal elements of Σ_t. Again, appealing to Campbell and Viceira's approximation method, the following log-linearization approximation holds:

$$\ln\left[R_f + \phi_{c,t}\left(R_{t+1}(\phi_{d,t}) + \alpha_t + \omega - R_f\right)\right] \\ \approx r_f + \phi_{c,t}\left[r_{\alpha,t+1}(\phi_{d,t}) - r_f\right] + \frac{1}{2}\phi_{c,t}(1-\phi_{c,t})\phi_{d,t}^T\Sigma_t\phi_{d,t}. \tag{55}$$

The optimal portfolio problem can be approximately rewritten as

$$\max_{\phi_{c,t}}(1-\gamma)^{-1}\ln\mathbb{E}_t\left\{e^{(1-\gamma)\left[\phi_{c,t}(\alpha_t+\omega)+\phi_{c,t}\phi_{d,t}^T(r_{t+1}-r_f\mathbf{1})+\frac{1}{2}\phi_{c,t}(1-\phi_{c,t})\phi_{d,t}^T\Sigma_t\phi_{d,t}+\frac{1}{2}\phi_{c,t}\phi_{d,t}^T\left(v_t-\Sigma_t\phi_{d,t}\right)\right]}\right\}.$$

After calculating the moment generating function and rearranging terms, searching for the optimal $\phi_{c,t}$ is equivalent to solving the following maximization problem:

$$\max_{\phi_{c,t}}\phi_{c,t}(\alpha_t+\omega) + \phi_{c,t}\phi_{d,t}^T\left(\mu_t - r_f + \frac{1}{2}v_t\right) - \frac{1}{2}\gamma\phi_{c,t}^2\phi_{d,t}^T\Sigma_t\phi_{d,t}. \tag{56}$$

The first-order condition is

$$0 = \alpha_t + \omega + \phi_{d,t}^T\left(\mu_t - r_f\mathbf{1} + \frac{1}{2}\nu_t\right) - \gamma\phi_{c,t}\phi_{d,t}^T\Sigma_t\phi_{d,t}. \tag{57}$$

Thus, according to Proposition 1.2, the optimal delegation $\phi_{c,t}$ is

$$\phi_{c,t} = \frac{1}{\gamma}\frac{1}{\phi_{d,t}^T\Sigma_t\phi_{d,t}}\left(\alpha_t + \omega + \gamma\phi_{d,t}^T\Sigma_t\phi_{d,t}\right) = 1 + \frac{\omega+\alpha_t}{\gamma_t}, \tag{58}$$

where the effective risk aversion is

$$\gamma_t \equiv S_t/\gamma, \quad \text{with } S_t \equiv \left(\mu_t - r_f\mathbf{1} + \frac{1}{2}\nu_t\right)^T\Sigma_t^{-1}\left(\mu_t - r_f\mathbf{1} + \frac{1}{2}\nu_t\right). \tag{59}$$

According to Proposition 1.1 and Equation (7), it holds that

$$\mu_t - r_f\mathbf{1} + \frac{1}{2}\nu_t = \left[(\rho L - I_n)\zeta_h + \frac{1}{2}\nu\right] h_t \text{ and } \Sigma_t = \Sigma h_t. \tag{60}$$

Therefore, by plugging (60) into (58), it follows that

$$\phi_{c,t} = 1 + \frac{\omega + \alpha_t}{\overline{\gamma} h_t}, \tag{61}$$

where $\overline{\gamma} = \left[(\rho L - I_n)\zeta_h + \frac{1}{2}\nu\right]^T \Sigma^{-1} \left[(\rho L - I_n)\zeta_h + \frac{1}{2}\nu\right] / \gamma$.

And hence, it holds that

$$q_t = \phi_{c,t}\frac{W_{c,t} - C_{c,t}}{W_t} \tag{62}$$

$$= \beta\lambda\left(1 + \frac{\omega + \alpha_t}{\overline{\gamma} h_t}\right). \tag{63}$$

Finally, after rearranging terms, it follows that

$$U_c(W_{c,t}) - U_d(W_{c,t}) = \beta\phi_c(\alpha_t + \omega) - \ln\left(1 - \frac{\overline{\alpha}}{\lambda} q_t\right) \tag{64}$$

$$= \beta\phi_c(\omega - \theta^{-1} q_t + \overline{\alpha} - f) - \ln\left(1 - \frac{\overline{\alpha}}{\lambda} q_t\right) \tag{65}$$

$$\geq \beta\phi_c(\omega - \theta^{-1} q_t + \overline{\alpha} - f) \tag{66}$$

$$\geq 0. \tag{67}$$

When $\omega + \overline{\alpha} > f + \theta^{-1}\lambda\beta$ as assumed in the proposition, the last inequality in (67) can be established by plugging in the equilibrium delegation q_t derived in Proposition 1.4.

2.4 Proof for Proposition 1.4

The equilibrium net alpha α_t and asset management services (i.e., delegation) q_t are determined by solving the intersection point of the following equations:

$$q_t = \theta(\overline{\alpha} - f) - \theta\alpha_t \quad (q_t \text{ supplied by funds}), \tag{68}$$

$$q_t = \beta\lambda \left[1 + (\omega + \alpha_t)/(\overline{\gamma}h_t)\right] \quad (q_t \text{ demanded by fund clients}). \tag{69}$$

Plugging (69) into (68) leads to the results.

2.5 Proof for Proposition 1.5

By definition, the aggregate fund flow is

$$\begin{aligned} flow_{t+1} &= \frac{Q_{t+1}}{Q_t} - R_{t+1}(\phi_{m,t}) - \alpha_t \\ &= \frac{q_{t+1}}{q_t}\frac{W_{t+1}}{W_t} - R_{t+1}(\phi_{m,t}) - \alpha_t \\ &= \frac{q_{t+1}}{q_t}\frac{W_{d,t} - C_{d,t} + (1-\overline{\alpha})Q_t}{W_t} R_{t+1}(\phi_t^{mkt}) - R_{t+1}(\phi_{m,t}) - \alpha_t. \end{aligned}$$

Thus, the aggregate fund flow can be rewritten as

$$flow_{t+1} = \frac{q_{t+1}}{q_t}\left[(1-\lambda)\beta + (1-\overline{\alpha})q_t\right] R_{t+1}(\phi_t^{mkt}) - R_{t+1}(\phi_{m,t}) - \alpha_t \tag{70}$$

$$= e^{\Delta \ln(q_{t+1}) + \ln[(1-\lambda)\beta + (1-\overline{\alpha})q_t] + r_{t+1}(\phi_t^{mkt})} - e^{r_{t+1}(\phi_{m,t})} - \alpha_t. \tag{71}$$

Log-linear approximation leads to

$$\begin{aligned} flow_{t+1} \approx\ & \Delta \ln(q_{t+1}) + \ln[(1-\lambda)\beta + (1-\overline{\alpha})q_t] \\ & + r_{t+1}(\phi_t^{mkt}) - r_{t+1}(\phi_{m,t}) - \alpha_t + \text{Jensen's term at } t. \end{aligned} \tag{72}$$

According to Proposition 1.1, it holds that

$$
\begin{aligned}
flow_{t+1} - \mathbb{E}_t\left[flow_{t+1}\right] &\approx \sqrt{h_t}\left[\frac{q'(\bar{h})}{q(\bar{h})}\sigma u_{t+1} + (\phi_t^{mkt} - \phi_{m,t})^T K u_{t+1} + (\phi_t^{mkt} - \phi_{m,t})^T \varepsilon_{t+1}\right] \\
&\approx \sqrt{h_t}\left[\frac{q'(\bar{h})}{q(\bar{h})}\sigma u_{t+1} + (\phi_t^{mkt} - \phi_{m,t})^T K u_{t+1}\right], \qquad (73)
\end{aligned}
$$

where the approximation in (73) is based on $(\phi_t^{mkt} - \phi_{m,t})^T \varepsilon_{t+1} \approx 0$ as n approaches infinity.

Given the market clearing condition on assets, we have the (approximated) relation in Theorem 2, which leads to

$$
\begin{aligned}
\phi_t^{mkt} &= \eta_t \phi_{m,t} + (1-\eta_t)\phi_{d,t} \\
&\approx \eta(\bar{h})\phi_{m,t} + [1-\eta(\bar{h})]\phi_{d,t},
\end{aligned}
$$

where $\eta_t \equiv q_t / \left[(1-\lambda)\beta + (1-\bar{\alpha})q_t\right]$.

Thus, it holds that

$$
\begin{aligned}
flow_{t+1} - \mathbb{E}_t\left[flow_{t+1}\right] &\approx \sqrt{h_t}\left\{\frac{q'(\bar{h})}{q(\bar{h})}\sigma u_{t+1} + [1-\eta(\bar{h})]\left(\phi_{d,t} - \phi_{m,t}\right)^T K u_{t+1}\right\} \\
&= \sqrt{h_t}\left\{\frac{q'(\bar{h})}{q(\bar{h})}\sigma u_{t+1} + [1-\eta(\bar{h})]\left(\Sigma_t^{-1}\mathcal{B}_t\right)^T K u_{t+1}\right\} \\
&= \sqrt{h_t}\left\{\frac{q'(\bar{h})}{q(\bar{h})}\sigma u_{t+1} + [1-\eta(\bar{h})]\mathcal{B}^T \Sigma^{-1} K u_{t+1}\right\} \\
&= \sqrt{h_t}\left\{\frac{q'(\bar{h})}{q(\bar{h})}\sigma u_{t+1} + [1-\eta(\bar{h})]\mathcal{B}^T \left(I_n + KK^T\right)^{-1} K u_{t+1}\right\}.
\end{aligned}
$$

According to Theorem 1, we can further obtain that

$$
flow_{t+1} - \mathbb{E}_t\left[flow_{t+1}\right] \approx \sqrt{h_t}\left\{\frac{q'(\bar{h})}{q(\bar{h})}\sigma u_{t+1} + [1-\eta(\bar{h})] A K^T \left(I_n + KK^T\right)^{-1} K u_{t+1}\right\}. \qquad (74)
$$

Therefore, the exposure coefficient is

$$
A = \frac{q'(\bar{h})}{q(\bar{h})}\sigma + [1-\eta(\bar{h})] A K^T \left(I_n + KK^T\right)^{-1} K. \qquad (75)
$$

2.6 Proof for Theorem 1

The portfolio choice is based on the competitive prices and aggregate fund flows in the equilibrium, including r_f, P_t, α_t, and $flow_{t+1}$. We can rewrite $R_{t+1}(\phi_{m,t}) + \alpha_t + flow_{t+1}$ as follows:

$$R_{t+1}(\phi_{m,t}) + \alpha_t + \pi_{t+1} = \tilde{R}_{t+1}(\tilde{\phi}_{m,t}) \tag{76}$$

$$\equiv R_f + \tilde{\phi}_m^T(\tilde{R}_{t+1} - R_f\mathbf{1}), \tag{77}$$

where

$$\tilde{\phi}_m \equiv \begin{bmatrix} 1 \\ \phi_m \end{bmatrix} \text{ and } \tilde{R}_{t+1} = \begin{bmatrix} R_f + \alpha_t + flow_{t+1} \\ R_{t+1} \end{bmatrix}. \tag{78}$$

Similar to Campbell and Viceira (1999, 2001), we can derive the approximation based on Proposition 1.5 as follows:

$$\ln(R_f + \alpha_t + flow_{t+1}) \approx \ln(1 + r_f + \alpha_t + flow_{t+1}) \tag{79}$$

$$\approx r_f + \alpha_t + flow_{t+1} - \frac{1}{2}AA^T h_t, \tag{80}$$

where $-\frac{1}{2}AA^T h_t$ is the Jensen's term. Therefore, the log returns are

$$\tilde{r}_{t+1} = \begin{bmatrix} r_f + \alpha_t - \frac{1}{2}AA^T h_t + flow_{t+1} \\ r_{t+1} \end{bmatrix}, \tag{81}$$

and the log returns are distributed as

$$\tilde{r}_{t+1} = \tilde{\mu}_t + \tilde{\Sigma}_t u_{t+1}, \tag{82}$$

where

$$\tilde{\mu}_t = \begin{bmatrix} r_f + \alpha_t - \frac{1}{2}AA^T h_t + \mathbb{E}_t\left[flow_{t+1}\right] \\ \mu_t \end{bmatrix} \text{ and } \tilde{\Sigma}_t = \begin{bmatrix} AA^T & AK^T \\ KA^T & \Sigma \end{bmatrix} h_t. \tag{83}$$

Now, we can apply the approximation of Campbell and Viceira (1999, 2001) again to obtain the following relation:

$$\tilde{r}_{t+1}(\tilde{\phi}_{m,t}) = \ln\left[\tilde{R}_{t+1}(\tilde{\phi}_{m,t})\right] \tag{84}$$

$$\approx r_f + \tilde{\phi}_{m,t}^T(\tilde{r}_{t+1} - r_f\mathbf{1}) + \frac{1}{2}\tilde{\phi}_{m,t}^T(\tilde{v}_t - \tilde{\Sigma}_t\tilde{\phi}_{m,t}), \tag{85}$$

where $\tilde{v}_t$ is the diagonal vector of $\tilde{\Sigma}_t$:

$$\tilde{v}_t = \begin{bmatrix} AA^T h_t \\ v_t \end{bmatrix}. \tag{86}$$

As a result, the augmented log returns are

$$\begin{aligned}\tilde{r}_{t+1}(\tilde{\phi}_{m,t}) &\approx r_f + (r_f + \alpha_t + flow_{t+1} - \frac{1}{2}AA^T h_t - r_f) + \phi_{m,t}^T(r_{t+1} - r_f\mathbf{1}) + \frac{1}{2}\tilde{\phi}_{m,t}^T\tilde{v}_t - \frac{1}{2}\tilde{\phi}_{m,t}^T\tilde{\Sigma}_t\tilde{\phi}_{m,t} \\ &= r_f + \alpha_t + flow_{t+1} - \frac{1}{2}AA^T h_t + \phi_{m,t}^T(r_{t+1} - r_f\mathbf{1} + \frac{1}{2}v_t) - \frac{1}{2}\phi_{m,t}^T\Sigma_t\phi_{m,t} - AK^T h_t\phi_{m,t}.\end{aligned}$$

Define $\mathcal{B}_t \equiv \mathcal{B}h_t$ with $\mathcal{B} = KA^T$. And thus, $\mathcal{B}_t$ is the covariance of the stock log returns and the aggregate flow:

$$\mathcal{B}_t = \mathrm{Cov}_t\left[r_{t+1}, flow_{t+1}\right]. \tag{87}$$

Then, the augmented log returns are

$$\tilde{r}_{t+1}(\tilde{\phi}_{m,t}) = r_f + \alpha_t + flow_{t+1} - \frac{1}{2}AA^T h_t + \phi_{m,t}^T(r_{t+1} - r_f\mathbf{1} + \frac{1}{2}v_t) - \frac{1}{2}\phi_{m,t}^T\Sigma_t\phi_{m,t} - \mathcal{B}_t^T\phi_{m,t}.$$

The optimal portfolio problem for fund managers can be simplified as

$$\max_{\phi_{m,t}}(1-\gamma)^{-1}\ln\mathbb{E}_t\left\{e^{(1-\gamma)\tilde{r}_{t+1}(\tilde{\phi}_{m,t})}\right\}. \tag{88}$$

After calculating the moment generating function, the optimal portfolio problem can be further rewritten as

$$\max_{\phi_{m,t}} \phi_{m,t}^T(\mu_t - r_f\mathbf{1} + \frac{1}{2}v_t - \mathcal{B}_t) - \frac{\gamma}{2}\phi_{m,t}^T\Sigma_t\phi_{m,t} + (1-\gamma)\phi_{m,t}^T\mathcal{B}_t. \tag{89}$$

The standard quadratic optimization problem leads to the optimal portfolio of fund managers:

$$\phi_{m,t} = \frac{1}{\gamma}\Sigma_t^{-1}\left(\mu_t - r_f + \frac{1}{2}\nu_t\right) - \Sigma_t^{-1}\mathcal{B}_t \tag{90}$$

$$= \phi_{d,t} - \Sigma_t^{-1}\mathcal{B}_t. \tag{91}$$

Because $\mathcal{B}_t = h_t\mathcal{B}$ and $\Sigma_t = h_t\Sigma$, it holds that

$$\phi_{m,t} = \phi_{d,t} - \Sigma^{-1}\mathcal{B}. \tag{92}$$

2.7 Proof for Corollary 1.2

The cross-sectional covariance between $\mathcal{B}_t$ and $\phi_{\tau,t}$ for each t is equal to

$$\text{Cov}\left[\mathcal{B}_t, \phi_{\tau,t}\right] = n^{-1}\mathcal{B}_t^T\Sigma_t^{-1}\mathcal{B}_t - n^{-2}\left(\mathbf{1}^T\mathcal{B}_t\right)\left(\mathbf{1}^T\Sigma_t^{-1}\mathcal{B}_t\right). \tag{93}$$

Because Σ_t is a positive definite symmetric matrix, according to the Cauchy-Schwarz inequality, it holds that

$$n^{-1}\mathcal{B}_t^T\mathbf{1}\mathbf{1}^T\Sigma_t^{-1}\mathcal{B}_t = n^{-1}(\mathcal{B}_t^T\mathbf{1}\mathbf{1}^T\Sigma_t^{-1/2})(\Sigma_t^{-1/2}\mathcal{B}_t) \tag{94}$$

$$\leq n^{-1}(\mathcal{B}_t^T\mathbf{1}\mathbf{1}^T\Sigma_t^{-1}\mathbf{1}\mathbf{1}^T\mathcal{B}_t)^{1/2}(\mathcal{B}_t^T\Sigma_t^{-1}\mathcal{B}_t)^{1/2}. \tag{95}$$

Thus, to show $\text{Cov}\left[\mathcal{B}_t, \phi_{\tau,t}\right] \geq 0$, it is sufficient to show that

$$n^{-1}\mathcal{B}_t^T\mathbf{1}\mathbf{1}^T\Sigma_t^{-1}\mathbf{1}\mathbf{1}^T\mathcal{B}_t \leq n^{-1}\mathcal{B}_t^T\Sigma_t^{-1}\mathcal{B}_t. \tag{96}$$

We denote $x \equiv n^{-1/2}\Sigma_t^{-1/2}\mathcal{B}_t$ and $y \equiv n^{-1/2}\Sigma_t^{-1/2}\mathbf{1}$, and thus, the inequality above can be rewritten as

$$x^T H_y x \leq x^T x, \tag{97}$$

where H_y is the orthogonal projection matrix, $H_y \equiv y(y^T y)^{-1} y^T$. Inequality (97) is obviously true once we realize that H_y is an orthogonal projection matrix.

2.8 Proof for Theorem 2

The market clearing condition of assets (ii.c) can be rewritten as

$$q_t \phi_{m,t} + \left[(1-\lambda)\beta - \overline{\alpha} q_t\right] \phi_{d,t} = \left[(1-\lambda)\beta + (1-\overline{\alpha}) q_t\right] \phi_t^{mkt}. \tag{98}$$

Plugging $\phi_{d,t} = \phi_{m,t} + \phi_{\tau,t}$ into the equation above, we obtain that

$$\phi_{m,t} = \phi_t^{mkt} - (1-\eta_t)\phi_{\tau,t}, \tag{99}$$

where $\eta_t \equiv q_t / \left[(1-\lambda)\beta + (1-\overline{\alpha}) q_t\right]$.

Therefore, the portfolio of direct investors is

$$\phi_{d,t} = \phi_t^{mkt} + \eta_t \phi_{\tau,t}. \tag{100}$$

2.9 Proof for Theorem 3

Based on the fund manager's optimal portfolio derived in Theorem 1 and the direct investor's optimal portfolio in Proposition 1.2, it holds that

$$\mu_t - r_f + \frac{1}{2}\nu_t = \gamma \Sigma_t \phi_{m,t} + \gamma \mathcal{B}_t. \tag{101}$$

According to the market clearing condition of assets (ii.c), it holds that

$$\phi_{m,t} = \eta_t^{-1}\phi_t^{mkt} - (\eta_t^{-1} - 1)\phi_{d,t} \tag{102}$$

$$= \eta_t^{-1}\phi_t^{mkt} - (\eta_t^{-1} - 1)\frac{1}{\gamma}\Sigma_t^{-1}(\mu_t - r_f + \frac{1}{2}v_t). \tag{103}$$

Plugging (103) into (101) and rearranging terms leads to

$$\mu_t - r_f\mathbf{1} + \frac{1}{2}v_t = \gamma\Sigma_t\phi_t^{mkt} + \eta_t\gamma\mathcal{B}_t. \tag{104}$$

Therefore, for any portfolio $r_{t+1}(\phi) = \phi^T r_{t+1}$ with $\mathbf{1}^T\phi = 1$, the risk premium is explained by the covariance with market return, denoted by r_{t+1}^{mkt}, and the covariance with common fund flow, denoted by $flow_{t+1}$:

$$\phi^T(\mu_t - r_f\mathbf{1} + \frac{1}{2}v_t) \approx \gamma\text{Cov}_t\left[r_{t+1}(\phi), r_{t+1}(\phi_t^{mkt})\right] + \eta_t\gamma\text{Cov}_t\left[r_{t+1}(\phi), flow_{t+1}\right]. \tag{105}$$

2.10 Proof for Corollary 1.3

According to Proposition 1.4 and Theorem 2, when $\lambda = 0$, $q_t = 0$ and thus $\eta_t = 0$. Therefore, Theorem 3 implies the conditional CAPM when $\lambda = 0$:

$$\phi^T(\mu_t - r_f\mathbf{1} + \frac{1}{2}v_t) \approx \gamma\text{Cov}_t\left[r_{t+1}(\phi), r_{t+1}(\phi_t^{mkt})\right] \tag{106}$$

$$= \gamma\text{Cov}_t\left[r_{t+1}(\phi), \hat{r}_{t+1}(\phi_t^{mkt})\right] \tag{107}$$

with $\hat{r}_{t+1}(\phi_t^{mkt}) \equiv r_{t+1}(\phi_t^{mkt}) - \mathbb{E}_t r_{t+1}(\phi_t^{mkt})$.

When $\lambda = 0$, the market portfolio is the mean-variance efficient portfolio:

$$\phi_t^{mkt} = \phi_{d,t} = \frac{1}{\gamma}\Sigma_t^{-1}(\mu_t - r_f\mathbf{1} + \frac{1}{2}\nu_t) \tag{108}$$

$$= \frac{1}{\gamma}\Sigma^{-1}\left[(\rho L - I_n)\zeta_h + \frac{1}{2}\nu\right]. \tag{109}$$

Thus, ϕ_t^{mkt} has constant portfolio weights, denoted by ϕ^{mkt}.

Further, according to (107), it holds that

$$\phi^T(\mu_t - r_f\mathbf{1} + \frac{1}{2}v_t) = \gamma Var_t\left[\hat{r}_{t+1}(\phi^{mkt})\right]\frac{\text{Cov}_t\left[r_{t+1}(\phi), \hat{r}_{t+1}(\phi^{mkt})\right]}{Var_t\left[\hat{r}_{t+1}(\phi^{mkt})\right]} \tag{110}$$

$$= \gamma Var_t\left[\hat{r}_{t+1}(\phi^{mkt})\right]\frac{\phi^T\Sigma\phi^{mkt}}{(\phi^{mkt})^T\Sigma\phi^{mkt}}. \tag{111}$$

Taking unconditional expectations on both sides leads to

$$\mathbb{E}\left[\phi^T r_{t+1} - r_f + \frac{1}{2}\phi^T v_t\right] = \Lambda\frac{\phi^T\Sigma\phi^{mkt}}{(\phi^{mkt})^T\Sigma\phi^{mkt}} \tag{112}$$

where $\Lambda \equiv \gamma\bar{h}\left[(\rho L - I_n)\zeta_h + \frac{1}{2}\nu\right]^T\Sigma^{-1}\left[(\rho L - I_n)\zeta_h + \frac{1}{2}\nu\right]$.

Lastly, $\frac{\phi^T\Sigma\phi^{mkt}}{(\phi^{mkt})^T\Sigma\phi^{mkt}}$ is actually the unconditional CAPM beta:

$$\frac{\phi^T\Sigma\phi^{mkt}}{(\phi^{mkt})^T\Sigma\phi^{mkt}} = \beta^{mkt}(\phi) \equiv \frac{\text{Cov}\left[r_{t+1}(\phi), \hat{r}_{t+1}(\phi^{mkt})\right]}{Var\left[\hat{r}_{t+1}(\phi^{mkt})\right]}. \tag{113}$$

Therefore, the unconditional CAPM holds:

$$\mathbb{E}\left[\phi^T r_{t+1} - r_f + \frac{1}{2}\phi^T v_t\right] = \beta^{mkt}(\phi)\Lambda. \tag{114}$$

References

Abel, Andrew B. 1987. "Operative gift and bequest motives." *American Economic Review*, 77(5): 1037–1047.

Ball, Ray, Gil Sadka, and Ronnie Sadka. 2009. "Aggregate earnings and asset prices." *Journal of Accounting Research*, 47(5): 1097–1133.

Bansal, Ravi, and Amir Yaron. 2004. "Risks for the long run: a potential resolution of asset pricing puzzles." *Journal of Finance*, 59(4): 1481–1509.

Barber, Brad M, Xing Huang, and Terrance Odean. 2016. "Which factors matter to investors? Evidence from mutual fund flows." *Review of Financial Studies*, 29(10): 2600–2642.

Barro, Robert J. 1974. "Are government bonds net wealth?" *Journal of Political Economy*, 82(6): 1095–1117.

Basak, Suleyman, and Anna Pavlova. 2013. "Asset prices and institutional investors." *American Economic Review*, 103(5): 1728–58.

Basak, Suleyman, Anna Pavlova, and Alexander Shapiro. 2007. "Optimal asset allocation and risk shifting in money management." *Review of Financial Studies*, 20(5): 1583–1621.

Ben-Rephael, Azi, Shmuel Kandel, and Avi Wohl. 2011. "The price pressure of aggregate mutual fund flows." *Journal of Financial and Quantitative Analysis*, 46(2): 585–603.

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

Berk, Jonathan B., and Jules H. van Binsbergen. 2016*a*. "Active managers are skilled: on average, they add more than $3 million per year." *Journal of Portfolio Management*, 42(2): 131–139.

Berk, Jonathan B, and Jules H van Binsbergen. 2016*b*. "Assessing asset pricing models using revealed preference." *Journal of Financial Economics*, 119(1): 1–23.

Berk, Jonathan B, and Richard C Green. 2004. "Mutual fund flows and performance in rational markets." *Journal of Political Economy*, 112(6): 1269–1295.

Brennan, Michael. 1993. "Agency and asset pricing." Anderson Graduate School of Management, UCLA University of California at Los Angeles, Anderson Graduate School of Management.

Caballero, Ricardo J., Emmanuel Farhi, and Pierre-Olivier Gourinchas. 2008. "An equilibrium model of "global imbalances" and low interest rates." *American Economic Review*, 98(1): 358–93.

Campbell, John Y., and Luis M. Viceira. 1999. "Consumption and portfolio decisions when expected returns are time varying." *Quarterly Journal of Economics*, 114(2): 433–495.

Campbell, John Y., and Luis M. Viceira. 2001. "Who should buy long-term bonds?" *American Economic Review*, 91(1): 99–127.

Cen, Xiao, Winston Wei Dou, Leonid Kogan, and Wei Wu. 2023. "Fund flows and income risk of fund managers." Working Paper.

Chapman, David A, Richard B Evans, and Zhe Xu. 2010. "The portfolio choices of young and old active mutual fund managers." Working Paper.

Cheng, Xu, Winston Wei Dou, and Zhipeng Liao. 2022. "Macro-finance decoupling: Robust evaluations of macro asset pricing models." *Econometrica*, 90(2): 685–713.

Chen, Hui, Winston Wei Dou, and Leonid Kogan. 2021. "Measuring the 'dark matter' in asset pricing models." *Journal of Finance*, Forthcoming.

Coval, Joshua, and Erik Stafford. 2007. "Asset fire sales (and purchases) in equity markets." *Journal of Financial Economics*, 86(2): 479–512.

Cuoco, Domenico, and Ron Kaniel. 2011. "Equilibrium prices in the presence of delegated portfolio management." *Journal of Financial Economics*, 101(2): 264–296.

Dou, Winston, and Adrien Verdelhan. 2017. "The volatility of international capital flows and foreign assets." University of Pennsylvania and MIT Working Paper.

Dou, Winston Wei, Leonid Kogan, and Wei Wu. 2023*a*. "Common fund flows: Flow hedging and factor pricing." *Journal of Finance*, Forthcoming.

Dou, Winston Wei, Leonid Kogan, and Wei Wu. 2023*b*. "Online appendix for "Common fund flows: Flow hedging and Factor Pricing"." *Journal of Finance*, Forthcoming.

Edelen, Roger M, and Jerold B Warner. 2001. "Aggregate price effects of institutional trading: a study of mutual fund flow and market returns." *Journal of Financial Economics*, 59(2): 195–220.

Edelen, Roger M., Richard B. Evans, and Gregory B. Kadlec. 2007. "Scale effects in mutual fund performance: The role of trading costs." Working Paper.

Edmans, Alex, Itay Goldstein, and Wei Jiang. 2012. "The real effects of financial markets: The impact of prices on takeovers." *Journal of Finance*, 67(3): 933–971.

Ferson, Wayne, and Jerchern Lin. 2014. "Alpha and performance measurement: The effects of investor disagreement and heterogeneity." *Journal of Finance*, 69(4): 1565–1596.

García, Diego, and Joel M. Vanden. 2009. "Information acquisition and mutual funds." *Journal of Economic Theory*, 144(5): 1965–1995.

Gennaioli, Nicola, Andrei Shleifer, and Robert Vishny. 2015. "Money doctors." *Journal of Finance*, 70(1): 91–114.

Gerakos, Joseph, Juhani T Linnainmaa, and Adair Morse. 2021. "Asset managers: Institutional performance and factor exposures." *Journal of Finance*, 76(4): 2035–2075.

Gómez, Juan-Pedro, and Fernando Zapatero. 2003. "Asset pricing implications of benchmarking: A two-factor CAPM." *European Journal of Finance*, 9(4): 343–357.

Gourinchas, Pierre-Olivier, and Hélène Rey. 2007. "International financial adjustment." *Journal of Political Economy*, 115(4): 665–703.

Grossman, Sanford J., and Joseph E. Stiglitz. 1980. "On the impossibility of informationally efficient markets." *American Economic Review*, 70(3): 393–408.

Hermalin, Benjamin E., and Michael S. Weisbach. 2012. "Information disclosure and corporate governance." *Journal of Finance*, 67(1): 195–233.

Hirshleifer, David, and Danling Jiang. 2010. "A financing-based misvaluation factor and the cross-section of expected returns." *Review of Financial Studies*, 23(9): 3401–3436.

Ibert, Markus, Ron Kaniel, Stijn Van Nieuwerburgh, and Roine Vestman. 2018. "Are mutual fund managers paid for investment skill?" *Review of Financial Studies*, 31(2): 715–772.

Kaniel, Ron, and Péter Kondor. 2013. "The delegated Lucas tree." *Review of Financial Studies*, 26(4): 929–984.

Khan, Mozaffar, Leonid Kogan, and George Serafeim. 2012. "Mutual fund trading pressure: Firm-level stock price impact and timing of SEOs." *Journal of Finance*, 67(4): 1371–1395.

Koijen, Ralph SJ. 2014. "The cross-section of managerial ability, incentives, and risk preferences." *Journal of Finance*, 69(3): 1051–1098.

Koijen, Ralph S. J., and Motohiro Yogo. 2019. "A demand system approach to asset pricing." *Journal of Political Economy*, 127(4): 1475–1515.

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh. 2018. "Interpreting factor models." *Journal of Finance*, 73(3): 1183–1223.

Leippold, Markus, and Roger Rueegg. 2020. "How rational and competitive is the market for mutual funds?" *Review of Finance*, 24(3): 579–613.

Long, J. Bradford De, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann. 1990. "Noise trader risk in financial markets." *Journal of Political Economy*, 98(4): 703–738.

Lou, Dong. 2012. "A flow-based explanation for return predictability." *Review of Financial Studies*, 25(12): 3457–3489.

Pástor, L'uboš, and Robert F Stambaugh. 2012. "On the size of the active management industry." *Journal of Political Economy*, 120(4): 740–781.

Pástor, L'uboš, Robert F Stambaugh, and Lucian A Taylor. 2015. "Scale and skill in active management." *Journal of Financial Economics*, 116(1): 23–45.

Pedersen, Lasse Heje. 2018. "Sharpening the arithmetic of active management." *Financial Analysts Journal*, 74(1): 21–36.

Perold, André F., and Robert S. Jr. Salomon. 1991. "The right amount of assets under management." *Financial Analysts Journal*, 47(3): 31–39.

Prat, Andrea. 2005. "The wrong kind of transparency." *American Economic Review*, 95(3): 862–877.

Ross, Stephen A. 1976. "The arbitrage theory of capital asset pricing." *Journal of Economic Theory*, 13(3): 341 – 360.

Roussanov, Nikolai, Hongxun Ruan, and Yanhao Wei. 2021. "Marketing mutual funds." *Review of Financial Studies*, 34(6): 3045–3094.

Sharpe, William F. 1991. "The arithmetic of active management." *Financial Analysts Journal*, 47(1): 7–9.

Shleifer, Andrei, and Robert Vishny. 1997. "The limits of arbitrage." *Journal of Finance*, 52(1): 35–55.

Stambaugh, Robert F., and Yu Yuan. 2016. "Mispricing factors." *Review of Financial Studies*, 30(4): 1270–1315.

Vayanos, Dimitri, and Paul Woolley. 2013. "An institutional theory of momentum and reversal." *Review of Financial Studies*, 26(5): 1087–1145.