Is There Too Much Benchmarking in Asset Management?

Anil K Kashyap*, Natalia Kovrijnykh†, Jian Li‡, and Anna Pavlova§

December 6, 2019

Abstract

The use of benchmarks for performance evaluation is commonplace in asset management, and yet, surprisingly, such contracts have not received much attention in the literature. This paper builds a model of delegated asset management in which benchmarking arises endogenously and analyzes the unintended consequences of benchmarking. The fund managers' portfolios are unobservable and so is the asset management cost. We show that conditioning managers' compensation on performance of a benchmark portfolio partially protects them from market risk and encourages them to generate more alpha. In general equilibrium, however, the use of such incentive contracts creates a pecuniary externality. Benchmarking inflates asset prices and gives rise to crowded trades, thereby reducing the effectiveness of incentive contracts for others. We show that privately-optimal contracts chosen by fund investors diverge from socially-optimal ones. A social planner, recognizing the crowding, opts for less benchmarking and less incentive provision. Privately-optimal contracts end up forcing managers to excessively pursue alpha, at too high a cost, and the planner corrects this. The planner's choice of benchmark portfolio weights also differs from the privately-optimal one.

JEL Codes: D82, D86, G11, G12, G23

Keywords: Asset Management, Portfolio Management, Benchmark, Index, Contracts, Incentives, Moral Hazard, Compensation, Welfare, Externality

We have benefited from comments and suggestions by seminar and conference participants at Arizona State University, Chicago Booth, London Business School, University of Oxford, UK Financial Conduct Authority, CERGE-EI Macro Workshop, 2019 SAET Conference, as well as discussions with Hector Chade and Dimitri Vayanos. This research has been supported by a grant from the Alfred P. Sloan Foundation to the Macro Financial Modeling project at the University of Chicago. The views expressed here are ours only and not necessarily those of the institutions with which we are affiliated.

*Booth School of Business, University of Chicago, National Bureau of Economic Research, Centre for Economic Policy Research and Bank of England. Email: anil.kashyap@chicagobooth.edu.

†Department of Economics, Arizona State University. Email: natalia.kovrijnykh@asu.edu.

‡Department of Economics, University of Chicago. Email: lijian@uchicago.edu.

§London Business School and Centre for Economic Policy Research. Email: apavlova@london.edu.

1 Introduction

Investors worldwide have delegated the investment of almost \$100 trillion to asset management firms. The portfolio managers' at these firms are invariably paid based on how their fund performs relative to a benchmark.[1] It is not immediately obvious why the compensation contracts take this form. In fact, there is little academic research on this question and there is no standard explanation for this phenomenon. We provide a theoretical framework that explains the common use of benchmarking in asset management. More importantly, we use it to assess the unintended general-equilibrium effects implications of benchmarking.

To study these questions, we embed an optimal-contracting model into a general-equilibrium setting. We show that optimally-designed contracts for fund managers involve benchmarking. This is because conditioning the managers' compensation on the performance of a benchmark portfolio partially protects them from risk, and thus gives them incentives to generate more alpha. In general equilibrium, however, the use of such incentive contracts creates a pecuniary externality through their effect on asset prices. Benchmarking inflates asset prices and reduces expected returns. This in turn reduces the marginal benefit of using incentive contracts for others. We show that a constrained social planner, who internalizes this externality, would opt for less incentive provision and less benchmarking.

Here is how our model works. Some agents in the economy, whom we call direct investors, manage their own money and others, whom we call fund investors, delegate their investment choice to others, whom we refer to as fund (or portfolio) managers. All agents are risk averse. The managers' portfolios are unobservable to fund investors and (part of) the cost of managing a portfolio is private. The managers are paid based on incentive contracts designed by the fund investors.[2] We focus on linear contracts which include a fixed salary, a fee for absolute performance, and potentially a fee for performance relative to a benchmark.

We assume that the managers can generate alpha through various sophisticated strategies. These include lending securities, conserving on transactions costs (e.g., from crossing trades in-house or by obtaining favorable quotes from brokers) or providing liquidity (i.e., serving as a counterparty to liquidity demanders and earning a premium on such trades). Such opportunities are not available to direct investors. While engaging in these activities

[1]For example, Ma, Tang, and Gómez (2019) report that around 80% of U.S. mutual funds explicitly base compensation on performance relative to a benchmark (usually a prospectus benchmark such as the S&P 500, Russell 2000, etc.).

[2]We abstract from the agency frictions between the fund investors and the asset-management firm and assume that the firm acts in best the interest of the fund investors, so that effectively the fund investors perfectly control the compensation arrangements for the portfolio managers. This is consistent with the fund trustees having a fiduciary obligation to their investors.

increases their alpha, portfolio managers have to incur a (parially) unobservable cost to implement them. The manager's portfolio is also unobservable/non-contractible. Thus fund investors design the manager's compensation contracts to incentivize the manager to produce alpha. A contract that pays the manager based on fund performance rewards her for taking these actions, but also exposes the manager to market risk. This risk, if unmitigated, means that the manager will underinvest in alpha-production. Adding a benchmark to the contract partially insulates the manager from market risk and strengthens her incentives to generate alpha.

We then turn to general equilibrium and analyze the unintended social consequences of benchmarking. When all fund investors use incentive contracts, they change the economy's aggregate demand for assets. In particular, benchmarking leads all managers to invest more in high-alpha assets and in assets that are in their benchmarks. The managers' demand boosts prices of such assets and lowers their expected returns. In other words, benchmarking contracts give rise to crowded trades.

Importantly, individual fund managers in our model take asset prices as given and they do not internalize the effects of contracts they design on equilibrium prices. Crowded trades resulting from the contract-induced incentives are in fact a pecuniary externality. Because of the agency frictions, this pecuniary externality leads to an inefficiency. Specifically, the use of benchmarking contracts by a group of investors reduces the effectiveness of contracts designed by other investors through crowded trades. This happens because asset prices enter the asset managers' incentive constraints. Each manager still has to incur the full private cost of managing assets but the benefits of doing so are reduced because of the crowded trades.

In light of this, a natural question to ask is how incentive contracts chosen by a social planner, who recognizes the effect of contracts on prices, differ from privately-optimal ones. We show that individual investors underestimate the cost of incentive provision relative to the social planner, who internalizes the negative externality of incentive contracts. As a result, the planner opts for less incentive provision. Specifically, we show that both the absolute-performance sensitivity as well as the level of benchmarking are lower in the socially-optimal contract than in the privately-optimal one. This ameliorates the price pressure that portfolio managers exert and reduces the crowdedness of trades.

Our model informs the debate as to whether costs of asset management are excessive and whether returns delivered by the managers justify these costs. We use the model to compare the managers' costs and expected returns under privately- and socially-optimal contracts. We find that, from the socially-optimal point of view, fund investors excessively rely on contracts and make their managers invest too much at too high a cost.[3] In the equilibrium with privately-optimal contracts, asset prices are higher and consequently expected returns are lower than those under socially-optimal contracts. Key to these implications is that, in contrast to fund investors, the social planner internalizes the pecuniary externality arising from crowded trades.

Finally, we investigate how benchmarks ought to be designed. We show that both privately- and socially-optimal benchmarks put more weight on assets for which portfolio management adds more value as well as on assets for which incentive misalignment is most serious. The relative tilt in the weights, however, is different in the privately- and socially-optimal benchmarks. For example, the planner puts relatively less weight on assets with large asset management costs compared to fund investors. This is because the planner understands that the cost of incentive provision is effectively higher than fund investors perceive it to be, and is therefore less willing to use benchmark weights for incentive provision.

The remainder of the paper is organized as follows. In the next section, we review the related literature. Section 3 presents our model, and Section 4 analyzes the model and derives our main results. Section 5 concludes and outlines suggestions for future research. Omitted proofs are in the Appendix.

2 Related Literature

Our work builds on the vast literature on optimal contracts under moral hazard, and in particular on seminal contributions of Holmstrom (1979) and Holmstrom and Milgrom (1987, 1991). Holmstrom (1979) argues that including in a contract a signal that is correlated with the output of the manager—in our case, such signal is the benchmark's performance—is beneficial to the principal. In our paper, the contract designer optimally chooses the signal to include in the contract. But more importantly, the benefit of including such signal is endogenous through the general-equilibrium effect on prices. To our knowledge, ours is the first paper that endogenizes the effectiveness of including an additional signal into an incentive contract.

Holmstrom and Milgrom (1991) introduce a tractable contracting setting with moral hazard, with which our model shares many similarities. The standard implication in this literature is that increasing the agent's share in the output of a project helps provide

[3]While the cost is borne by the manager, it ultimately gets passed on to the fund investor, who needs to compensate the manager enough to ensure her participation.

incentives to the agent. In the context of delegated asset management though, giving the agent a larger share of portfolio return encourages her to scale down the risk of the (unobservable) portfolio by reducing risky asset holdings. Stoughton (1993) and Admati and Pfleiderer (1997) show that the manager is able to completely "undo" her steeper incentives by such scaling, and her incentives to collect information on asset payoffs remain unchanged. In our paper, we design a contract that provides desired incentives, despite the endogenous portfolio response of the manager, and show that it involves benchmarking. Another notable difference from the aforementioned literature is that we embed optimal contracts in a general-equilibrium setting and study interactions between contracts and equilibrium prices, and the implications of these interactions on welfare.

Our work is also related to the literature in asset pricing and corporate finance theory that explores the general-equilibrium implications of benchmarking. The pioneering work of Brennan (1993) shows that benchmarking leads to lower expected returns on stocks included in the benchmark. Cuoco and Kaniel (2011) and Basak and Pavlova (2013) study benchmarking in dynamic models, and show that the positive price pressure on benchmark stocks pushes up their prices and lowers their Sharpe ratios. Basak and Pavlova also show that benchmarking leads to excess volatility and excess co-movement of returns on stocks inside the benchmark. Kashyap, Kovrijnykh, Li, and Pavlova (2018) focus on implications of benchmarking portfolio managers on firm's corporate decisions and demonstrate that firms in the benchmark have a higher valuation for investment projects or merger targets than firms outside the benchmark. All this literature takes the benchmarking contract of managers to be exogenous. The only exceptions are Buffa, Vayanos, and Woolley (2014) and Cvitanic and Xing (2018), who study asset-pricing implications of benchmarking in an environment with endogenous contracts.[4] In both models, benchmarking helps reduce diversion of cash flows from the fund by managers. Our rationale for benchmarking is to reward alpha-generating activities.

Our paper also relates to the literature on pecuniary externalities in competitive equilibrium settings with incomplete markets, for example, Lorenzoni (2008), He and Kondor (2016), Gromb and Vayanos (2002), Davila and Korinek (2018), Biais, Heider, and Ho-

[4]See also Ozdenoren and Yuan (2017) who conduct a related analysis in the context of an industry equilibrium, in a classical moral-hazard setting with many principal-agent pairs. They show that benchmarking is privately-optimal but it creates overinvestment and excessive risk-taking at the industry level. Albuquerque, Cabral, and Guedes (forthcoming) present a related model of industry equilibrium, enriched further with strategic interactions among firms in the industry, and show that benchmarking against peer performance induces agents to take correlated actions.

erova (2019), and Acemoglu and Simsek (2012), among others.[5] Lorenzoni (2008) studies a model of credit booms, where a pecuniary externality arises from the combination of limited commitment and asset prices being determined in spot markets. Decentralized equilibria feature over-borrowing relative to what is constrained optimal (although there is always under-borrowing compared to the first best). Borrowing less ex ante is welfare improving as it leads to an increase in asset prices in the low-output state, which allows entrepreneurs to transfer resources to the low-output state. Both our setting and mechanism are very different, but we share a similar prediction that asset prices in decentralized equilibrium fall between those in constrained and unconstrained optima. However, the actual comparison of prices is reverse in our model—decentralized-equilibrium prices are higher than in the constrained optimum (lower in his paper), but lower than in the first best (higher in his paper).

He and Kondor (2016) study a model where individual firms' liquidity management decisions generate investment waves. These investment waves are constrained inefficient when future investment opportunities are noncontractible, and the social and private value of liquidity differs. In their model, overinvestment occurs during booms and underinvestment occurs during recessions. Gromb and Vayanos (2002) analyze a model where competitive financially-constrained arbitrageurs supply liquidity to the market, and fail to internalize that changing their positions affects prices. A change in prices effectively moves resources across time and states and thus can bring the marginal rates of substitution closer together and be Pareto improving. A social planner can achieve a Pareto improvement by either reducing or increasing the arbitrageurs' position. Davila and Korinek (2018) highlight a distinction between "distributive externalities" that arise from incomplete insurance markets and "collateral externalities" that arise from price-dependent financial constraints. The externality that we emphasize in our paper falls into the second category, broadly defined, although in our case the inefficiency arises from incentive rather than financial constraints.

Biais, Heider, and Hoerova (2019) analyze a model, where protection buyers trade derivatives with protection sellers, and there is moral hazard on the side of protection sellers. In their model, although competitive equilibrium prices enter incentive constraints, a pecuniary externality does not lead to constrained inefficiency as it does in our model. The reason is that in their setup investors optimally supply insurance against the risk of fire sales. In Acemoglu and Simsek (2012), firms trade off providing insurance to workers and incentivizing them to exert effort. The authors show that, under certain conditions, equi-

[5] This literature goes back to Hart (1975), Greenwald and Stiglitz (1986), and Geanakoplos and Polemarchakis (1996).

librium prices can tighten incentive constraints. They mainly focus on inefficient sharing of idiosyncratic risk. Instead, our focus is on the inefficient use of an additional signal—return of the benchmark portfolio— in the incentive contract.

Finally, Fershtman and Judd (1987) look at contract design in an equilibrium setting, but in an oligopoly rather than a competitive equilibrium, as in our model. In their paper, there is strategic manipulation of agents' incentive, as owners take into account that the contracts they give to their managers affect contracts chosen by other owners. In our competitive setting (in a very different environment), private agents ignore their effects on others.

3 Model

We embed a linear optimal-contracting problem into a general-equilibrium asset-pricing framework. Each fund investor designs a compensation contract for his manager; in doing so, he takes equilibrium asset prices as given. In equilibrium, managers' demands, governed by their compensation contracts, end up affecting equilibrium prices. This, in turn, affects the effectiveness of individual contracts in the economy, as well as social welfare. The focus of our analysis is on the optimal compensation contracts and their equilibrium effects on prices and welfare.

3.1 Economy

Except for portfolio managers and their clients, our environment is standard. There are two periods, $t = 0, 1$. Investment opportunities are represented by N risky assets (stocks) and one risk-free bond. The risky assets are claims to cash flows $\tilde{D}$, realized at $t = 1$, where $\tilde{D} \sim N(\mu, \Sigma)$. The variables $\tilde{D}$ and μ are $N \times 1$ vectors and Σ is an invertible $N \times N$ matrix. The risk-free bond pays an interest rate that is normalized to zero. The risky assets are in a fixed supply of $\bar{x} > 0$ shares, where $\bar{x}$ is an $N \times 1$ vector. The bond is in infinite net supply. Let S, an $N \times 1$ vector, denote asset prices. Asset prices are determined endogenously in equilibrium.

There is a continuum of agents in the economy, of three types. First, there are "direct" investors—constituting a fraction λ_D of the population—who manage their own portfolios. There are also portfolio or fund managers—a fraction λ_M—and fund investors who hire those managers—a fraction λ_F, with $\lambda_D + \lambda_M + \lambda_F = 1$. We further assume for simplicity that each fund investor employs one manager, so that $\lambda_M = \lambda_F$.[6] Fund investors can buy the bond directly, but cannot trade risky assets, so they delegate the selection of their portfolios to managers. Each agent has a constant absolute risk aversion (CARA) utility function over final wealth (or compensation) W, $U(W) = -e^{-\gamma W}$, where $\gamma > 0$ is the coefficient of absolute risk aversion. Fund investors and direct investors are endowed with x_{-1}^F and x_{-1}^D shares of risky assets, respectively, where $\lambda_F x_{-1}^F + \lambda_D x_{-1}^D = \bar{x}$.

For fund investors, delegating investment to a portfolio manager will have costs and benefits. On the one hand, as we will discuss in the next subsection, managers can potentially outperform direct investors. On the other hand, the fund investor cannot dictate to the manager what portfolio she should choose. That is, the manager's portfolio choice is not contractible, or, equivalently, unobservable. It is true that some managers, e.g., mutual funds, are required to disclose their portfolios at a particular point in time. However, their actual portfolios between the disclosure dates differ significantly from their reported portfolios (Kacperczyk, Sialm, and Zheng, 2008), and a fund investor cannot obtain detailed information on the manager's trades. So the unobservability assumption is very realistic.

Furthermore, the managers incur a private cost of managing a portfolio. The combination of the private cost and the portfolio choice being unobservable will be the central friction in our model.

We do not model an agent's choice to become a direct investor or a fund investor—the fractions of different investors in the population are exogenous. One could endogenize this choice, for example, by assuming heterogeneous costs of participating in the asset market. We, however, abstract from this as it is not central to the main message of the paper.

3.2 Value Added and Costs of Asset Management

We assume that portfolio managers can potentially outperform direct investors. The (per-share) return for a direct investor's portfolio x is given by $x^\top(\tilde{D} - S)$. The manager's returns are

$$r_x = x^\top(\Delta + \tilde{D} - S) + \varepsilon, \tag{1}$$

where $\Delta > 0$ is an (exogenous) vector and $\varepsilon \sim N(0, \sigma_\varepsilon)$ is a (scalar) noise term. The manager incurs a private portfolio-management cost $\psi^\top x$, where $\psi > 0$ is an exogenous

[6]The extension where one manager is hired by multiple investors acting collectively is straightforward.

vector.[7] So managers in our model incur costs to generate excess returns of $x^\top \Delta$ that we call α. The fact that this cost is private and the portfolio choice x is unobservable, will be the key driving force of our results.[8]

In this formulation, the managers' α has nothing to do with superior information, which gives rise to superior stock-selection and market-timing abilities. If it did, then any direct investors who happened to buy the same assets or traded at the same time, without any knowledge of the Δ's, would earn the same returns. So this setup is consistent with the vast literature (e.g., Fama and French, 2010) that casts doubt on the ability to generate abnormal returns by stock picking or market timing.

Instead, the managers' α comes from activities such as lending securities, delivering lower transactions costs (e.g., from crossing trades in-house or by obtaining favorable quotes from brokers) or providing liquidity (i.e., serving as a counterparty to liquidity demanders and earning a premium on such trades). We refer to these activities as "return-augmenting" strategies.

A critical assumption for our results is that the manager bears a private cost in delivering the abnormal returns. The existence of this cost is most clear-cut for the provision of liquidity. To successfully buy and sell at the appropriate times, the manager has to be actively monitoring market conditions while markets are open. For securities lending, the manager will also have to decide whether to accommodate requests to borrow shares. In some cases, these demands arise because the entity borrowing the shares wants to vote them and the manager must decide whether to pass up that choice. Adams, Mansi, and Nishikawa (2014) find that outsourcing the lending process to a third party yields lower returns than arranging this in-house. It seems reasonable that not all of these costs can be

[7] Our assumption regarding managers' α-generating activities in (1) warrants further elaboration. Implicit in our expressions for the returns on the fund and the portfolio-management cost is that they scale linearly with the size of the portfolio. This seems to be inconsistent with Berk and Green (2004) who assume that there are decreasing returns to scale in asset management. There is no inconsistency. Berk and Green explicitly attribute decreasing returns to scale to the price impact of fund managers. The bigger the portfolio invested in an α-opportunity, the smaller the return on a marginal dollar invested. Berk and Green's model is in partial equilibrium and their price impact is simply an exogenous function of fund size. Ours is a general equilibrium model, in which the price impact endogenously arises from a higher aggregate demand of portfolio managers, leading to lower expected returns of high-Δ assets. The linearity of returns and costs in the size of the portfolio allows us to solve the model in closed form. We conjecture that our main results go through as long as the cost is strictly increasing in x. We are planning to incorporate this analysis in the next draft of the paper.

[8] One alternative formulation of the private cost that we have investigated assumes the manager needs to exert effort that is privately costly (and unobservable) to shareholder to generate the excess returns. That setup, which is often employed in the contracting literature (e.g. Holmstrom and Milgrom, 1987, 1991), yields a similar set of implications, though the algebra is more tedious. This equivalence gives us further confidence that the reduced-form choice that we have made is reasonable.

recouped. For the crossing of trades, we are less certain of how the internal decisions about this choice is made. Funds also report that client service requests tend to scale with the size of the portfolio and these costs are not fully recouped.[9]

The most powerful and convincing justification for the assumption that there are some private costs is that, if we accept the idea that some abnormal returns are obtainable, then it follows that you might naturally expect all managers to reap similar gains. Yet, the foregoing evidence suggests that there are major differences in the extent to which different managers pursue different strategies and thus earn different levels of revenue. If there were not some costs involved, then this heterogeneity would be very puzzling.

Turning to the revenues, there is a number of empirical papers that establish the profitability of securities lending, trading cost minimization and liquidity provision. We briefly review some of this evidence, explaining as we go why direct investors would not be able to mimic portfolio managers to earn the same returns.

It is well-documented that securities lending is profitable. For example, it contributed 5% of total revenue of both BlackRock and State Street in 2017. Dimensional Fund Advisors' prospectuses report the net revenues from securities lending as a percentage of a portfolio's average daily net assets. There is tremendous variation across their different funds. For instance in 2018, the U.S. Large Company Portfolio earned 1 basis point, while the U.S. Small Cap Portfolio earned 11 basis points and the Emerging Markets Small Cap Portfolio earned 78 basis points. While it has recently become possible for some retail investors to participate in securities lending, they earn lower returns for this activity and do not have the same opportunities as a large asset management firm.

It is also well-established that portfolio managers can profit from providing immediacy in trades, by either buying assets which are out of favor or selling ones that are in high demand. In a classic paper, Keim (1999) estimates an annual α of 2.2% earned by liquidity provision activities of a fund. Rinne and Suominen (2016) document that mutual funds differ noticeably in the extent to which they either demand or supply liquidity. They estimate that the top decile of supplying funds outperform the bottom decile by about 60 basis points per year. Anand, Jotikasthira, and Venkataraman (2018) find similar estimates using a different sample of funds over a different time period. It would be prohibitively expensive for retail investors to try to do this.

[9] Most managers also incur some costs that are observable and can be passed on directly to fund investors. Examples would include custody, audit, shareholder reports, proxies and some external legal fees. The analysis of the model where some costs are observable is available from the authors upon request and will be incorporated in the next draft of the paper.

While practitioners often mention the advantages that portfolio managers have in minimizing transactions costs, there is less empirical work quantifying these gains. One exception is Eisele, Nefedova, Parise, and Peijnenburg (forthcoming), which draws on transaction-level data that allows them to compare trades inside a fund complex with equivalent ones between outside parties. By crossing buy and sell orders, the affiliated transaction can save the bid-ask spread that would normally be paid. They find that indeed since the Securities and Exchange Commission began monitoring the within complex deals in 2004, the crossed trades are executed more cheaply than comparable external trades. Again, direct investors simply do not have the ability to follow this strategy.

Another modeling assumption is that the aforementioned benefits are proportional to the size of the holdings. For the liquidity provision and trade-crossing, this assumption strikes us as obvious. The wider the range of securities in the portfolio and/or the more a fund holds on any particular security, the easier it will be to provide liquidity or more likely it will be that a trade can be offset. For securities lending, by expanding the range of assets in the portfolio we would expect to open up additional lending opportunities. Whether holding more of any particular asset improves lending opportunities is less obvious. The demand for borrowing stocks or bonds varies greatly over time and across securities, so knowing whether having a large position will be valuable is uncertain.

For simplicity, we assume that Δ (the expected per-share return coming from the return-augmenting activities) is exogenous. One might argue that as more managers engage in, e.g., securities lending, the lower is the fee that receive for it due to general-equilibrium considerations. Our mechanism would work through an endogenous Δ the same way as it works through the endogenous S—fund investors, taking the (per-share) returns $\Delta + \tilde{D} - S$ as given, ignore the fact that their contracts push down the equilibrium returns.[10] Thus endogenizing Δ would not change (but even potentially strengthen) our results.

Finally, the noise term ε in (1) captures the fact that aforementioned return-augmenting activities do not produce a certain return each period. For example, the demand for liquidity, the opportunities to lend share and the possibility of crossing-trades all fluctuate so even a very alert and skilled manager will have some randomness in their returns. Also for securities that are lent, there is a risk that they will not be returned in a timely manner or potentially at all.[11]

[10] A way of explicitly modeling the market for the return-augmented activities would depend on which exact activity is considered. Since we attempt to capture several of such activities, we abstract from fully modeling such a market.

[11] One might wonder what happens of the noise is proportional to x (that is, the noise term is εx instead of ε). The algebra is messier in this case, but the main results go through. The analysis of this case is

3.3 Managers' Compensation Contracts

To provide incentives for the managers to invest into the risky assets and to generate α, the fund investors design compensation contracts. The managers receive compensation w from fund investors. This compensation has three parts: one is a linear payout based on absolute performance of the manager's portfolio x, the second part depends on the performance relative to the benchmark portfolio, and the third is independent of performance.[12] Then

$$w = \hat{a} r_x + b(r_x - r_{\mathbf{b}}) + c = a r_x - b r_{\mathbf{b}} + c, \tag{2}$$

where r_x is the performance of the manager's portfolio defined in (1), $r_{\mathbf{b}} = \theta^\top(\tilde{D} - S)$ is the performance of the benchmark portfolio θ, and $a = \hat{a} + b$ (all of our analysis and the intuitions that follow will be in terms of a rather than $\hat{a}$). The contract is (a, b, c, θ)—or, equivalently, $(\hat{a}, b, c, \theta)$—where a is sensitivity to absolute performance, or "skin-in-the-game", b is sensitivity to relative performance, c is the fixed component, and θ is the $N \times 1$ vector of benchmark weights. The contract for a particular manager is optimally chosen by the fund investor who employs her.

The restriction of linearity of optimal contracts warrants a discussion. We assume linearity for the purpose of tractability.[13] Characterizing fully-optimal contracts is hard in general. It is even harder in our case since we solve for them in a general-equilibrium model where contracts affect equilibrium prices and thus in turn affect the contract chosen by each fund investor. The restriction to linearity allows us to find optimal contracts in closed form. However, the argument behind our main mechanism—individual contracts change demand functions, which in the aggregate affects equilibrium prices, which dampens the effect of contracts on demands, and thus makes contracts less effective—is very general. We have no reason to believe that this mechanism would not apply with fully-general contracts.

We think of a manager's contract as a compensation contract between a portfolio manager and her investment-advisor firm (e.g., BlackRock, who we assume is acting in the

available from the authors upon request.

[12]This part captures features such as a fee linked to initial assets under management or a fixed salary or any other fixed costs

[13]There is a literature that justifies the use of linear contracts in environments with CARA preferences and normally-distributed returns. Holmstrom and Milgrom (1987) show that, in a specific dynamic setting, the solution of the optimal-contracting problem is as if the problem were the static one and the principal is constrained to use a linear compensation rule that depends on the final outcome. Holmstrom and Milgrom restrict the agent's action to affect only the mean of the random process but not the variance, which is not the case in our model (the portfolio choice affects both the mean and the variance of the return). Sung (1995) establishes the robustness of the Holmstrom and Milgrom's linearity result by allowing the agent to also control the variance.

interests of the fund investors). For U.S. mutual funds, we have a great deal of information about such contracts. Since 2005 mutual funds in the U.S. have been required to include a "Statement of Additional Information" in the prospectus that describes how portfolio managers are compensated. Ma, Tang, and Gómez (2019) analyze this information and find that around 80% of the funds explicitly base compensation on performance relative to a benchmark (usually the prospectus benchmark, e.g., S&P 500, Russell 2000, etc.). Some funds state in their Statements of Additional Information that managers are encouraged to invest in their own funds, i.e., have skin-in-the-game. Khorana, Servaes, and Wedge (2007) document mutual fund managers' holdings of their own funds, which provides evidence for an absolute performance component in their compensation. Ma, Tang, and Gómez report that most managers also have a fixed salary component, but the fraction of fund managers whose entire compensation consists of only fixed salary is very small. The performance-based bonus exceeds the fixed salary for 68% of the funds in the sample, constituting more than 200% of fixed salary for 35% of funds.[14] Evidence on compensation contracts of portfolio managers other than mutual funds is sparse because these managers are not required to report their compensation structure and they rarely do so voluntarily. Bank for International Settlements (2003) presents survey-based evidence for a sample of other fund managers including sovereign wealth funds and pension funds, and also finds evidence supporting our compensation structure, and in particular finds that performance evaluation relative to benchmarks is pervasive.

4 Analysis

4.1 Direct Investors' and Managers' Problems

At $t = 0$, direct investors choose a portfolio of risky assets x and the risk-free bond holdings to maximize their expected utility $-Ee^{-\gamma W}$. Since their return on the portfolio is $x^\top(\tilde{D}-S)$, the resulting time-1 wealth is $W = \left(x_{-1}^D\right)^\top S + x^\top(\tilde{D} - S)$. It is well-known that a direct investor's maximization problem is equivalent to the following mean-variance optimization:

$$\max_x x^\top(\mu - S) - \frac{\gamma}{2} x^\top \Sigma x. \tag{3}$$

Portfolio managers choose a portfolio of risky assets x and the risk-free bond holdings to

[14] In contrast, Ibert, Kaniel, Van Nieuwerburgh, and Vestman (2017) find surprisingly weak sensitivity of manager pay to performance for Swedish mutual funds.

maximize $-E\exp\{-\gamma[ar_x - br_{\mathbf{b}} + c - x^\top \psi]\}$, where the quantity inside the square brackets is their compensation net of private cost. This maximization problem is equivalent to the following optimization:

$$\max_x \quad ax^\top(\Delta - \psi/a + \mu - S) + b\theta^\top(\mu - S) + c - \frac{\gamma}{2}\left[(ax - b\theta)^\top\Sigma(ax - b\theta) + a^2\sigma_\varepsilon^2\right], \quad (4)$$

where we have substituted r_x and $r_{\mathbf{b}}$ defined above. The first three terms are the manager's expected pay, net of her private cost, and the last term is the variance of her compensation, scaled by a half of her absolute risk aversion γ. One important observation we make at this stage is related to the first term in (4): the manager receives a fraction a of the per-share abnormal return on the assets, Δ, but pays the entire cost ψ per share. (We later show that $a < 1$.)

Both the direct investors and managers take asset prices as given. Lemma 1 reports the optimal portfolios of the direct investors and managers arising from their optimizations.

Lemma 1 (**Portfolio Choice**). *The direct investors' and managers' portfolio demands are as follows:*

$$x^D = \Sigma^{-1}\,\frac{\mu - S}{\gamma}, \quad (5)$$

$$x^M = \Sigma^{-1}\,\frac{\Delta - \psi/a + \mu - S}{a\gamma} + \frac{b\theta}{a}. \quad (6)$$

A direct investor's portfolio is the standard mean-variance portfolio, scaled by his risk aversion γ. A manager's portfolio differs from that of a direct investor in three dimensions. First, managers split their risky assets investments between two portfolios: the (modified) mean-variance portfolio and the benchmark portfolio. The latter arises because the manager's compensation is exposed to fluctuations in the benchmark. To mitigate this risk, she holds a hedging portfolio that is (perfectly) correlated with the benchmark, i.e., the benchmark itself. This implication is very general, and we share it with other models that analyzed benchmarking, both in two-period and multi-period economies and for other investor preferences specifications.[15] The split between the two portfolios is governed by the strength of the relative-performance incentives, captured by b. The higher b is, the closer the manager's portfolio is to the benchmark.

Second, because our managers have access to return-augmenting strategies, they per-

[15]This result first appears in Brennan (1993) in a two-period model. Cuoco and Kaniel (2011) and Basak and Pavlova (2013), among others, obtain it in dynamic models with different preferences.

ceive the mean-variance tradeoff differently from the direct investors and tilt their mean-variance portfolios towards high-Δ assets. Consistent with this result, Johnson and Weitzner (2019) report that portfolio managers' portfolios in their sample overweight assets with high securities-lending fees.

Finally, the manager scales her (modified) mean-variance portfolio by $1/a$ relative to that of a direct investor. The reason for the scaling is that, as we can see from the first term in (2), for each share that an manager holds, she gets a fraction a of the total return.

Let us now focus on the variance of the manager's compensation, $(ax^M - b\theta)^\top \Sigma (ax^M - b\theta) + a^2 \sigma_\varepsilon^2$, which we have first encountered in (4), and compare two managers, one of whom does not have a benchmark ($b = 0$) and one who does ($b > 0$). Absent benchmarking, the safest investment choice for the manager is to invest everything in the riskless asset ($x^M = 0$). With benchmarking, the safest investment choice is to hold the benchmark portfolio ($x^M = b\theta/a$), or to be a "closet indexer." Both managers would like to take some risk (and earn the corresponding return), and they do so by investing in the mean-variance portfolio. For the same skin-in-the-game a, the benchmarked manager is willing to hold $b\theta/a$ *more shares* in the risky assets. The variances of the compensations of the two managers are exactly the same.[16] This is the sense in which benchmarking partially protects the manager from risk.

We are now ready to solve for equilibrium prices. Substituting portfolio demands from Lemma 1 into the market-clearing condition for assets, $\lambda_M x^M + \lambda_D x^D = \bar{x}$, we find

$$S = \mu - \gamma \Sigma \Lambda \bar{x} + \gamma \Sigma \Lambda \lambda_M \frac{b\theta}{a} + \Lambda \frac{\lambda_M}{a} \left(\Delta - \frac{\psi}{a} \right), \tag{7}$$

where

$$\Lambda \equiv \left[\frac{\lambda_M}{a} + \lambda_D \right]^{-1} \tag{8}$$

modifies the market's effective risk aversion.

To develop intuition, it is useful to compare the asset prices in our economy, (7), with their counterparts in the economy with direct investors only ($\lambda_M = 0,\ \lambda_D = 1$). In that economy $S = \mu - \gamma \Sigma \bar{x}$, a familiar expression. We can see that the managers, who have an additional demand for benchmark assets and for high-Δ assets, push prices of such assets up (relative to otherwise identical low-benchmark-weight and low-Δ assets). Specifically, the term $\gamma \Sigma \Lambda \lambda_M b\theta/a$ in (7) accounts for the price pressure due to benchmarking and the term $\Lambda(\lambda_M/a)(\Delta - \psi/a)$ for α-chasing. In our economy, higher prices imply lower expected

[16] To see this formally, note that $ax^M - b\theta = \Sigma^{-1}(\Delta - \psi/a + \mu - S)/\gamma$ does not depend on b.

returns. Therefore, in the aggregate, the managers exert price impact and lower expected returns on the assets that they particularly desire to hold in their portfolios.[17]

4.2 Fund investors' Problem

Each fund investor chooses the contract (a, b, c, θ) and portfolio $x = x^M$ to maximize his expected utility

$$-E \exp\left\{-\gamma\left[\left(x^F_{-1}\right)^\top S + r_x - (ar_x - br_{\mathbf{b}}) - c\right]\right\}$$

subject to the manager's participation constraint

$$-E \exp\left\{-\gamma\left[(ar_x - br_{\mathbf{b}}) + c - x^\top \psi\right]\right\} \geq u_0, \tag{9}$$

and her incentive constraint (6). The latter is the first-order condition of the manager's optimization problem, capturing the fact that the portfolio x is the manager's private choice. The right-hand side of the participation constraint, u_0, is the value of manager's outside option.[18]

Equivalently, we can rewrite the fund investor's problem in terms of mean-variance utilities. It will be convenient to express payoffs in terms of the following variable. Denote $y = ax^M - b\theta$ and $z = x^M - y$, which are effective allocations of asset holdings to the manager and fund investor, respectively. Furthermore, define the fund investor's and manager's mean-variance utilities as

$$U^F\left(a, \frac{b\theta}{a}, y, S\right) = \left(x^M\right)^\top (1-a)\Delta + z^\top(\mu - S) - \frac{\gamma}{2}\left[z^\top \Sigma z + (1-a)^2\sigma^2_\varepsilon\right] + \left(x^F_{-1}\right)^\top S - c,$$

$$U^M(a, y, S) = \left(x^M\right)^\top (a\,\Delta - \psi) + y^\top(\mu - S) - \frac{\gamma}{2}\left[y^\top \Sigma y + a^2\sigma^2_\epsilon\right] + c,$$

where

$$x^M = \frac{y}{a} + \frac{b\theta}{a}, \tag{10}$$

$$z = x^M - y = \frac{1-a}{a}y + \frac{b\theta}{a}. \tag{11}$$

[17] Note that this discrepancy cannot be arbitraged away in our model. Direct investors are free to engage in any arbitrage activity investors because the are unrestricted in their portfolio choice. As long as managers represent a meaningful fraction of the market (i.e., λ_M is non-negligible), however, there are always differences in prices of assets that depend on their benchmark weights and managers' Δ.

[18] We do not model explicitly what this outside option is, as it does not matter for our main results. It can be exogenous, or it can be endogenized.

Then the fund investor's problem is to maximize his utility subject to the manager's participation constraint, and her (modified) incentive constraint,

$$\max_{a,b,c,\theta,y} \quad U^F \tag{12}$$

$$\text{s.t. } U^M \geq \tilde{u}_0, \tag{13}$$

$$y = \Sigma^{-1}\frac{\Delta - \psi/a + \mu - S}{\gamma}, \tag{14}$$

where the value $\tilde{u}_0$ is the mean-variance version of u_0.[19]

Notice that because of the contract's constant component c, in the mean-variance formulation utility becomes transferable, and the fund investor effectively maximizes the total utility of the fund investor and manager subject to the manager's incentive constraint. The manager's participation constraint is then trivially satisfied by adjusting the constant c.

We next discuss the roles that the contract parameters a, b, and θ play in the fund investor's maximization problem.

4.3 Contracts and Incentives

As a point of reference, consider the first best where the manager's portfolio choice is observable and contractible. The first-best contract involves perfect risk sharing between the (equally risk-averse) fund investor and manager and no benchmarking, $a = 1/2$ and $b = 0$.[20] It is easy to show that if the manager were facing the first-best contract but chose the portfolio privately, she would overestimate the cost of α-production, thus investing less in all assets, especially assets with high ψ. A higher a reduces the manager's effective cost ψ/a, which increases her demand for risky assets, especially those with a high cost of α-production. However, a higher a also exposes the manager to more risk, which makes her scale down x^M, as reflected in the denominator(s) of (10). Thus the use of absolute performance creates a tension between incentive provision and risk sharing.

The use of benchmarking, together with an appropriate benchmark selection, alleviates this tension by mitigating the second, adverse effect of a. As we have discussed in Section 4.1, benchmarking shields the manager from risk by reducing variance in her compensation *for a given portfolio choice.*[21] As a result, (for the same a) the manager invest

[19]In particular, if the manager's outside option is risk-free, then $u_0 = -\exp(-\gamma\tilde{u}_0)$.

[20]See Lemma 5 in the Appendix for the formal analysis.

[21]By reducing in the manager's risk exposure, benchmarking makes it cheaper for the fund investor to implement any particular portfolio choice.

more. In addition, if the benchmark portfolio puts a relatively higher weight on certain assets, the manager's exposure to risk is reduced more for those assets, and she will invest proportionally more in them. That is, benchmarking protects the manager from risk, and an appropriate choice of the benchmark portfolio can help to improve incentives for α-production.

4.4 Privately-Optimal Contracts

Notice that the fund investor fully internalizes the manager's cost of managing the fund. Formally, this can be seen by taking the first-order condition with respect to c, which immediately implies that the Lagrange multiplier on the participation constraint equals one. But since the manager bears the cost privately and only receives fraction a of the return, for her the effective cost is higher, which is why ψ/a appears in (6). These two costs being different plays an important role in our analysis.

All the main tradeoffs will be apparent from the first-order conditions. It will also be useful to compare the first-order conditions of a fund investor to those of a social planner. Therefore in what follows, we present and carefully discuss those conditions.

Notice that b enters into the fund investor's and manager's problems only though $b\theta/a$. Therefore we take the first-order condition with respect to $b\theta/a$, and later derive the expression for b separately. The first-order condition with respect to $b\theta/a$ is given by

$$\frac{\partial(U^F + U^M)}{\partial(b\theta/a)} = \Delta - \psi + \mu - S - \gamma\Sigma z = 0. \tag{15}$$

It captures the marginal effect on the total utility of the fund investor and manager due to a higher demand by the managers in response to more benchmarking. Substituting (11) and (14) into the above equation and rearranging terms (see the Appendix), we have

$$\gamma\Sigma b\theta = (2a-1)(\Delta - \psi + \mu - S) + (1-a)\left(\frac{1}{a} - 1\right)\psi. \tag{16}$$

The two terms on the right hand side of equation (16) capture two concerns that fund investors have in mind when designing the benchmark. Consider two special cases, the first one is $a = 1/2$ when perfect risk-sharing is achieved, and the second one is $a = 1$ when the private and social costs are aligned. As we will show later, in the optimal contract $a \in (1/2, 1)$, so both terms on the right-hand side of (16) are positive. The first term, $(2a-1)(\Delta - \psi + \mu - S)$, is there because the fund investor recognizes that benchmarking increases the total expected surplus net of cost. Since $a > 1/2$, the manager is exposed to more risk than unconstrained optimal, so the fund investor uses benchmarking to make her invest more, in particular in assets with a higher value added $\Delta - \psi$. The second term, $(1-a)(1/a-1)\psi$, reflects the incentive-provision role of $b\theta$. By protecting the manager from risk, benchmarking provides her with incentives to invest more. Such incentive provision is especially important for assets with high cost, ψ, as the manager is the most reluctant to invest in them.

Let us now examine the first-order condition with respect to a, which is given by[22]

$$\begin{aligned} 0 &= \frac{\partial(U^F + U^M)}{\partial a} + \frac{\partial U^F}{\partial y}\frac{\partial y}{\partial a} \\ &= -(2a-1)\gamma\sigma_\epsilon^2 - (\Delta - \psi + \mu - S - \gamma\Sigma z)^\top \frac{y}{a^2} + \frac{1-a}{a}(\Delta + \mu - S - \gamma\Sigma z)^\top \frac{\partial y}{\partial a} \\ &= -(2a-1)\gamma\sigma_\epsilon^2 + \frac{1-a}{a}\psi^\top \frac{\partial y}{\partial a}, \end{aligned} \tag{17}$$

where the last equality follows from (15). Notice the appearance of $\partial y/\partial a$ in the above equation. It captures how a marginal increase in a affects the surplus through the direct response of the manager's (modified) demand for the risky assets, y. This is the incentive-provision channel. The other terms reflect that the design of a also affects how risk is split between the fund investor and the manager (and thus also how much of the risky asset the manager buys). This is the risk-sharing channel.

Using (15) and $\partial y/\partial a = \Sigma^{-1}\psi/(\gamma a^2)$ (obtained by differentiating (14) with respect to a), the above equation becomes

$$(1-a)\frac{\psi^\top \Sigma^{-1}\psi}{\gamma a^3} - (2a-1)\gamma\sigma_\varepsilon^2 = 0. \tag{18}$$

Notice again terms multiplying $(1-a)$ and $(2a-1)$ appear, except unlike in (16), they now have different signs. This means that there is a tradeoff that these two terms represent. A higher a is beneficial as it provides incentives for α-production, which is captured by the first term, but is also costly as it exposes the manager to too much risk, as captured by the second term.

Substituting the expression for S, we obtain closed-form expressions for equilibrium contracts given in part (a) of the next lemma.

Lemma 2. *In the equilibrium with the privately-optimal contract,*

[22]The manager's utility is maximized with respect to y, so $(\partial U^M/\partial y)(\partial y/\partial a)$ does not appear in (17).

(a) $a = a^{private}$, $b = b^{private}$, *and* $\theta = \theta^{private}$ *are given by*

$$0 = (1-a)\frac{\psi^\top \Sigma^{-1}\psi}{\gamma a^3} - (2a-1)\gamma\sigma_\varepsilon^2, \tag{19}$$

$$b = (2a-1)\,\mathbf{1}^\top \left[\bar{x} + \frac{\Sigma^{-1}}{\gamma}\lambda_D(\Delta-\psi)\right] + (1-a)\left[\frac{1}{a} - \left(\frac{\lambda_M}{a} + \lambda_D\right)\right]\mathbf{1}^\top \frac{\Sigma^{-1}}{\gamma}\psi, \tag{20}$$

$$\theta = \frac{2a-1}{b}\left[\bar{x} + \frac{\Sigma^{-1}}{\gamma}\lambda_D(\Delta-\psi)\right] + \frac{1-a}{b}\left[\frac{1}{a} - \left(\frac{\lambda_M}{a} + \lambda_D\right)\right]\frac{\Sigma^{-1}}{\gamma}\psi; \tag{21}$$

(b) *asset prices are given by*

$$S^{private} = \mu - \gamma\Sigma\bar{x} + \lambda_M\left(\Delta - \psi + \Delta - \frac{\psi}{a}\right) \tag{22}$$

and the managers' asset holdings are

$$x^M_{private} = 2\bar{x} + \frac{\Sigma^{-1}}{\gamma}\lambda_D\left(\Delta - \psi + \Delta - \frac{\psi}{a}\right). \tag{23}$$

Notice that (19) is the expression in a only. Then given a, (20) delivers the expression for b, and finally, given a and b, (21) gives us expression for the benchmark weights θ. We will use (20) to provide sufficient conditions for the fund investor to use benchmarking in what follows. Before we proceed with that, let us briefly comment on the expression for the equilibrium prices given by (22). Recall that in the economy without managers the equilibrium prices would be $S = \mu - \gamma\Sigma\bar{x}$. Prices are higher in the presence of managers due to their higher demands as engage in return-augmenting activities, as captured by the last term in (22). Notice that it contains $\Delta - \psi$ and $\Delta - \psi/a$, which are the extra expected returns net of costs as perceived by the fund investors and by the managers, respectively. Similarly, the equilibrium asset holdings of managers in (23) are higher when the opportunities for α-production are better. Notice that managers hold exactly $2\bar{x}$ when $\lambda_D = 0$. We will discuss this special case further in subsection 4.5.

For some of our results on benchmarking, we will need to impose the following assumption.

Assumption 1. *Suppose that*

(a) $\mathbf{1}^\top\left[\bar{x} + \lambda_D\Sigma^{-1}(\Delta-\psi)/\gamma\right] > 0$,

(b) $\mathbf{1}^\top\Sigma^{-1}\psi > 0$.

Assumption 1 is a mild technical restriction. It is trivially satisfied when the variance-covariance matrix Σ is diagonal or when Δ's and ψ's are the same for all assets (and given that $\Delta - \psi \geq 0$). When Σ is not diagonal (which implies that cross-price elasticities of the manager's demand function are not zero), it is useful to interpret the assumption as follows. Part (a) is a necessary and sufficient condition for the sum of shares (over all assets) that the manager holds in the first best to be positive (that is, the manager does not borrow).[23] Part (b) means that if the private cost ψ increases by the same percentage for all assets, then the sum of shares (over all assets) that the manager holds in equilibrium goes down. In other words, the manager reduces total holdings when the cost is higher.

Using the above assumption and the equilibrium expression for b presented in Lemma 2, we have the following result:

Proposition 1 (**Benchmarking is Optimal**). *Suppose that Assumption 1 holds. Then the privately-optimal contract involves benchmarking, that is,* $b^{private} > 0$.

Proposition 1 helps us understand why benchmarking in the asset management industry is so pervasive. Benchmarking helps fund investors because it incentivizes the manager to engage more in risky α-generating activities by partially protecting her from some risks. In the language of the asset management industry, benchmarked managers are being protected from "β" (i.e., the fluctuations in the return of the benchmark portfolio) while being rewarded for α.

Proposition 1 is essentially a version of Holmstrom (1979) famous sufficient-statistic result—the use of an additional signal (in this case, the benchmark return) helps the contract designer provide incentives to the manager in a more effective way. Holmstrom's result is general and does not require contract linearity, so while we have restricted our attention to linear contracts, we would expect that the manager's compensation would depend on the benchmark return even with more general contracts.

Next, we discuss the properties of the privately-optimal benchmark weights. Using equation (21), the lemma below shows how these weights differ across assets with different value added or cost of α-production, which are $\Delta - \psi$ and ψ respectively.

Lemma 3. *Consider two assets, i and j, that have the exact same characteristics except $\Delta_i - \psi_i \geq \Delta_j - \psi_j$ and $\psi_i \geq \psi_j$, with at least one inequality being strict. Then in the privately-optimal contract, asset i has a larger weight in the benchmark than asset j:* $\theta_i^{private} > \theta_j^{private}$.

[23] See the proof of Lemma 5 in the Appendix.

The reason for this result is intuitive: fund investors recognize that manipulating benchmark weights allows them to provide more incentives for investment in assets where α-production is the most valuable. The effect of a larger ψ on the benchmark weight is ambiguous, as can be seen from (21). On the one hand, the incentive problem is the most severe for assets with a larger ψ, and thus setting higher weight is most valuable for those assets. On the other hand, a larger ψ reduces the total expected return, which reduces the marginal benefit of using $b\theta$ for protecting the manager from extra risk. However, for the same (or a larger) value added, higher-cost assets would have a higher weight in the privately-optimal benchmark.

Fund investors design contracts to influence the manager's demand for risky assets. Through the aggregate demand of the managers, contracts influence equilibrium asset prices, as given by (7). Prices then affect the marginal cost/marginal benefit tradeoff of contracts for individual fund investors. Since fund investors take prices as given, they do not internalize how their choices of contracts (once aggregated) affect effectiveness of other fund investors' contracts. In other words, fund investors impose an externality on each other through their use of contracts. It is natural to wonder then what the contract would be if chosen by a social planner, who is subject to the same restrictions as fund investors, but internalizes the effect of contracts on prices. We explore this issue in the next subsection.

4.5 Socially-Optimal Contracts

We define the problem of such a constrained social planner as follows. The planner maximizes the weighted average of fund investors' and direct investors' utilities subject to the participation and incentive constraints of the managers, as well as the constraint that direct investors choose their portfolios themselves.[24] As before this problem can be equivalently rewritten in terms of mean-variance preferences.[25] Specifically, define

$$U^D = \left(x_{-1}^D\right)^\top S + \left(x^D\right)^\top (\mu - S) - \frac{\gamma}{2}\left(x^D\right)^\top \Sigma x^D.$$

[24]Equivalently, instead of imposing the manager's participation constraint, her utility can be included into the planner's objective function with a Pareto weight ω_M. For the transfer c to be finite, we must have $\omega_M = \omega_F$. This is analogous to noticing that the Lagrange multiplier on the participation constraint, which effectively acts as the Pareto weight on the manager, equals ω_F.

[25]We provide the original formulation in terms of exponential utilities in the Appendix.

Then the social planner's problem is

$$\max_{a,b,c,\theta,y} \;\; \omega_F U^F + \omega_D U^D$$

subject to (13), (14), and (5).

The social planner's first-order condition with respect to $b\theta/a$ is

$$\begin{aligned} 0 = &\left[\omega_F \left(x^F_{-1} - x^M\right)^\top + \omega_D \left(x^D_{-1} - x^D\right)^\top\right] \frac{\partial S}{\partial(b\theta/a)} \\ &+ \omega_F \left[\frac{\partial(U^F + U^M)}{\partial(b\theta/a)} + \frac{\partial U^F}{\partial y}\frac{\partial y}{\partial S}\frac{\partial S}{\partial(b\theta/a)}\right]. \end{aligned} \tag{24}$$

The terms in the first line of the above equation capture what we call the redistribution effect. Depending on the initial endowments and the Pareto weights, the social planner has incentives to use benchmarking to move prices so as to benefit one or the other party based on this redistribution motive. We discuss the redistribution effects in Remark 1 at the end of this subsection. To isolate the effects of benchmarking not coming from this redistribution motive, which we find more interesting as they capture the actual externality, we set the Pareto weights $\omega_F = \omega_D$. Then by market clearing, $\omega_F \left(x^F_{-1} - x^M\right) + \omega_D \left(x^D_{-1} - x^D\right) = 0$, the term in the first line of (24) is zero.[26] Rewriting the term in the second line, (24) becomes

$$\begin{aligned} 0 &= (\Delta - \psi + \mu - S - \gamma\Sigma z)^\top + \frac{1-a}{a}(\Delta + \mu - S - \gamma\Sigma z)^\top \frac{\partial y}{\partial S}\frac{\partial S}{\partial(b\theta/a)} \\ &= (\Delta - \psi + \mu - S - \gamma\Sigma z)^\top - \frac{1-a}{a}(\Delta + \mu - S - \gamma\Sigma z)^\top \frac{\lambda_M}{\lambda_M/a + \lambda_D}, \end{aligned} \tag{25}$$

or, equivalently,

$$\Delta - \psi\frac{\lambda_M/a + \lambda_D}{\lambda_M + \lambda_D} + \mu - S - \gamma\Sigma z = 0. \tag{26}$$

Compare (25) or (26) with (15). The second term in (either line of) (25) captures the general-equilibrium externality that the planner is trying to correct, and it is negative. The planner realizes that benchmarking inflates prices and thus reduces returns. Hence for the social planner the benefit of α-production is smaller due to this crowded-trades effect, or,

[26] Choosing Pareto weights to cancel out the redistribution effects is equivalent to allowing the social planner to use transfers. The planner will then use transfers to equate the marginal utilities (weighted by Pareto weights) of different agents.

equivalently, the cost is higher for the same unit of benefit: $\psi(\lambda_M/a + \lambda_D)/(\lambda_M + \lambda_D) > \psi$ in (26) instead of ψ in (15). So from the social planner's point of view, α-chasing is less beneficial/more expensive, which, as we will see, will make her do less of it in equilibrium.

Substituting the expression for z,

$$\gamma\Sigma b\theta = (2a-1)\left(\Delta - \frac{\lambda_M/a + \lambda_D}{\lambda_M + \lambda_D}\psi + \mu - S\right) + (1-a)\left(\frac{1}{a} - \frac{\lambda_M/a + \lambda_D}{\lambda_M + \lambda_D}\right)\psi. \tag{27}$$

Again, compared to (16), the social cost is inflated.

The planner's first-order condition with respect to a (after again canceling out the redistribution effects) is

$$\begin{aligned} 0 &= \frac{\partial(U^F + U^M)}{\partial a} + \frac{\partial U^F}{\partial y}\left[\frac{\partial y}{\partial a} + \frac{\partial y}{\partial S}\frac{\partial S}{\partial a}\right] \\ &= (1-2a)\gamma\sigma_\epsilon^2 - (\Delta - \psi + \mu - S - \gamma\Sigma z)^\top \frac{y}{a^2} + \frac{1-a}{a}(\Delta + \mu - S - \gamma\Sigma z)^\top \left[\frac{\partial y}{\partial a} + \frac{\partial y}{\partial S}\frac{\partial S}{\partial a}\right]. \end{aligned}$$

Comparing this to (17), there is an additional term containing $(\partial y/\partial S)(\partial S/\partial a)$. It reflects that the planner understands that contracts affect prices, which in turn affect the managers' demands and thus the marginal benefit of α-production. However, unlike in the first-order condition with respect to $b\theta/a$, we cannot sign this extra term—recall that the effect of a on the manager's incentives is ambiguous. That is, *for a given* $b\theta/a$, the social planner's benefit of using a can be higher or lower than that of an individual fund investor. What is interesting is that once the planner takes into account the adjustment in the social optimal $b\theta$, the effect of a that reduces x^M and thus lowers prices is exactly offset by this adjustment. Hence the additional term that remains in the first-order condition with respect to a is only the part that takes into account how a higher a increases incentives for investment, which in turn increases prices and reduces returns. As a result, the marginal benefit of a for the social planner is lower than that for individual fund investors, and the possibility of benchmarking is a crucial for this result.

To show this formally, use (25) to rewrite the above equation as follows:[27]

$$
\begin{aligned}
0 &= -(2a-1)\gamma\sigma_\epsilon^2 + \frac{1-a}{a}\frac{\lambda_M/a+\lambda_D}{\lambda_M+\lambda_D}\psi^\top\left[\frac{\partial y}{\partial a} + \frac{\partial y}{\partial S}\frac{\partial S}{\partial a} + \frac{y}{a^2}\frac{\partial y}{\partial S}\frac{\partial S}{\partial(b\theta/a)}\right] \\
&= -(2a-1)\gamma\sigma_\epsilon^2 + \frac{1-a}{a}\frac{\lambda_M/a+\lambda_D}{\lambda_M+\lambda_D}\psi^\top\left[\frac{\partial y}{\partial a} - \frac{\partial y}{\partial a}\frac{\lambda_M/a}{\lambda_M/a+\lambda_D}\right].
\end{aligned}
$$

We can see that the effectiveness of incentive provision for the planner, captured by the term proportional to $\partial y/\partial a$, is smaller than for private fund investors in equation (18). Finally, using $\partial y/\partial a = \Sigma^{-1}\psi/(\gamma a^2)$, we obtain the analog of (18),

$$
(1-a)\frac{\psi^\top\Sigma^{-1}\psi}{\gamma a^3}\frac{\lambda_D}{\lambda_M+\lambda_D} - (2a-1)\gamma\sigma_\varepsilon^2 = 0. \tag{28}
$$

The benefit of incentive provision captured by the first term is smaller than the corresponding term in (18). Comparing (28) with (18), it is easy to see that the social planner will use a lower a than individual fund investors. We will formalize this result later in Proposition 2.

Substituting for the equilibrium prices, the following lemma described the equilibrium contract and prices in closed form.

Lemma 4. *In the equilibrium with the socially-optimal contract,*

(a) $a = a^{social}$, $b = b^{social}$ *and* $\theta = \theta^{social}$ *are given by*

$$
0 = (1-a)\frac{\psi^\top\Sigma^{-1}\psi}{\gamma a^3}\frac{\lambda_D}{\lambda_M+\lambda_D} - (2a-1)\gamma\sigma_\varepsilon^2, \tag{29}
$$

$$
b = (2a-1)\mathbf{1}^\top\left[\bar{x} + \frac{\Sigma^{-1}}{\gamma}\lambda_D(\Delta-\psi)\right] + (1-a)\left[\frac{1}{a} - \frac{\lambda_M/a+\lambda_D}{\lambda_M+\lambda_D}\right]\mathbf{1}^\top\frac{\Sigma^{-1}}{\gamma}\psi, \tag{30}
$$

$$
\theta = \frac{2a-1}{b}\left[\bar{x} + \frac{\Sigma^{-1}}{\gamma}\lambda_D(\Delta-\psi)\right] + \frac{1-a}{b}\left[\frac{1}{a} - \frac{\lambda_M/a+\lambda_D}{\lambda_M+\lambda_D}\right]\frac{\Sigma^{-1}}{\gamma}\psi; \tag{31}
$$

(b) asset prices are given by

$$
S^{social} = \mu - \gamma\Sigma\bar{x} + \lambda_M\left(\Delta - \frac{\lambda_M/a+\lambda_D}{\lambda_M+\lambda_D}\psi + \Delta - \frac{\psi}{a}\right) \tag{32}
$$

[27] To get the second line, differentiate the market-clearing condition $\lambda_M(y/a + b\theta/a) + \lambda_D x^D = 0$ with respect to $b\theta/a$ and a and use $\partial x^D/\partial S = \partial y/\partial S$ to get $\left(\frac{\lambda_M}{a}+\lambda_D\right)\frac{\partial y}{\partial S}\frac{\partial S}{\partial(b\theta/a)} + \lambda_M = 0$ and $\left(\frac{\lambda_M}{a}+\lambda_D\right)\frac{\partial y}{\partial S}\frac{\partial S}{\partial a} - \lambda_M\frac{y}{a^2} + \frac{\lambda_M}{a}\frac{\partial y}{\partial a} = 0$ so that $\left(\frac{\lambda_M}{a}+\lambda_D\right)\left[\frac{\partial y}{\partial S}\frac{\partial S}{\partial a} + \frac{y}{a^2}\frac{\partial y}{\partial S}\frac{\partial S}{\partial(b\theta/a)}\right] + \frac{\lambda_M}{a}\frac{\partial y}{\partial a} = 0$.

and the managers' asset holdings are

$$x^M_{social} = 2\bar{x} + \frac{\Sigma^{-1}}{\gamma}\lambda_D \left(\Delta - \frac{\lambda_M/a + \lambda_D}{\lambda_M + \lambda_D}\psi + \Delta - \frac{\psi}{a}\right). \tag{33}$$

Equations (29)–(31) are the analogs of (19)–(21). As expected, the two sets of equations coincide when $\lambda_M = 0$, and hence there is no externality. But so long as there are managers, the socially- and privately-optimal contracts are different. Our analysis below will reveal how exactly they compare to each other.

Notice how the equilibrium price in (32) compares to that in (22). The fact that the cost as perceived by the planner is higher than that perceived by an individual fund investor is reflected in the price, and reduces the price for a given a. This reflects the fact that the planner internalizes adverse effect of incentive provision and views it as being more costly compared to fund investors. This causes the planner to provide less incentives, which in particular means that $a^{social} < a^{private}$, as we will show in Proposition 2. Lower value of a further reduces the price, and thus $S^{social} < S^{private}$, as we will see in Proposition 3.

We are now ready to present the central result of the paper.

Proposition 2 (Socially- vs. Privately-Optimal Contracts). *(a) Compared to the equilibrium with the privately-optimal contract, the socially-optimal contract has a smaller absolute-performance sensitivity, that is,* $a^{social} < a^{private}$*;*

(b) Suppose that Assumption 1 holds. Then compared to the equilibrium with the privately-optimal contract, the socially-optimal contract involves less benchmarking, that is, $b^{social} < b^{private}$.[28]

As we have seen in our analysis, the use of incentive contracts inflates prices and thus *reduces the marginal benefit of incentive provision* for everyone else. The social planner internalizes this effect, and opts for less incentive provision than individual fund investors.

An interesting special case is the one when there are no direct investors, i.e., $\lambda_D = 0$. Notice that in this case each manager will hold exactly $2\bar{x}$ and the total α in the economy is fixed, equal to $2\bar{x}^\top\Delta$. The planner understands that incentive provision is unnecessary in this case, and thus faces no tradeoff between incentive provision and incentives. Indeed, by substituting $\lambda_D = 0$ into (29)–(30), it immediately follows that the socially-optimal contract is $a = 1/2$ and $b = 0$, which provides perfect risk sharing and coincides with the first-best one (see Lemma 5 in the Appendix). Interestingly, the fund investors ignore the

[28] It is also true that $b^{social}/a^{social} < b^{private}/a^{private}$.

fact that, in equilibrium, their contracts will not help them generate higher returns, and use contracts with $a > 1/2$ and $b > 0$, as can be seen from (19)−(20).

To further emphasize that benchmarking is crucial for the comparison between privately- and socially-optimal contracts, consider a case where benchmarking is exogenously set to zero, $b = 0$. In this case, all incentive provision and risk sharing has to be done through a. As we discussed earlier, an increase in a has two opposing effects on the managers' demands and hence prices. As a result, it can be shown that with $b = 0$ the comparison between a^{social} and $a^{private}$ is ambiguous. Intuitively, both the marginal benefit of a (incentive provision) as well as its marginal cost (exposing the manager to more risk) are lower for the social planner who internalizes the effect of a on prices. Depending on which reduction is bigger, the planner can choose a higher or a lower a than the fund investors do. Thus, only because of benchmarking ($b \neq 0$) can we be sure of the direction of the externality and are able to say that privately-optimal contracts deliver excessive incentive provision.

We now show that excessive incentive provision and excessive benchmarking give rise to crowded trades and excessive costs.

Proposition 3 (Crowded Trades and Excessive Costs of Asset Management). *Compared to the equilibrium corresponding to the privately-optimal contract, in the equilibrium corresponding to the socially-optimal contract*

(a) asset prices are lower, $S^{social} < S^{private}$;

(b) managers' costs are lower, $\psi^\top x^M_{social} < \psi^\top x^M_{private}$.

As Proposition 3 shows, excessive use of incentive contracts by fund investors inflates prices and reduces per-share returns. In addition, the costs of asset management are excessively high. Our model thus contributes to the debate on whether costs of asset management are excessive and whether returns delivered by the managers justify these costs.

Finally, we discuss the benchmark weights. The same statement as in Lemma 3 applies to the case with socially-optimal contracts. In addition, we can compare the tilts to high value-added and/or high-cost assets in the privately- and socially-optimal contracts.

Claim 1 (Comparison of Benchmark Weights). *Suppose that Assumption 1 holds. Then privately-optimal benchmark underweights assets with higher $\Delta - \psi$ and overweighs assets with higher ψ compared to the socially-optimal benchmark. Formally, consider two assets i and j, that have the exact same characteristics except $\Delta_i - \psi_i \geq \Delta_j - \psi_j$ and $\psi_i \leq \psi_j$, with at least one inequality being strict. Then $\theta_i^{social} - \theta_j^{social} > \theta_i^{private} - \theta_j^{private}$.*

The intuition behind this result is a little tricky. Compare (21) and (31), and recall that the role of $b\theta$ is to protect the manager from risk as well as provide incentives, as captured by $(2a-1)$- and $(1-a)$-terms, respectively. The planner understands that the second role is less effective relative to how fund investors perceive it. That role is driven by ψ, and hence the planner is more reluctant to use benchmark weights to provide incentives for high-ψ assets. And as the role of $b\theta$ in protecting the manager from risk is relatively more important than incentive provision, she will tilt the benchmark more towards high-value-added assets than individual fund investors would.

Remark 1 (**Redistribution Effects).** Through our choice of weights in the social-welfare function, we have shut down the contracts' redistribution effects and isolated the pecuniary externality that the planner desires to correct. For certain applications (e.g., concerned with wealth inequality), however, it is useful to highlight the transfers from one set of agents to another that benchmarking generates. The answer is simple and it depends on agents' initial endowments or, more precisely, on whether an agent is a (net) buyer of assets or a (net) seller. As we have argued, benchmarking boosts asset prices. This benefits (net) sellers of the assets at the expense of (net) buyers. If the social planner favors investors who have high endowments of assets and are planning to sell (e.g., the older generations), she has incentives to use more benchmarking in order to inflate prices so as to benefit them, and vice versa if she favors net buyers (who are typically the younger generations).

Remark 2 (**Prices Relative to the First Best).** According to Proposition 3, $S^{social} < S^{private}$. Surprisingly, the asset prices in the first-best case exceed equilibrium prices under both privately- and socially-optimal contracts, that is, $S^{social} < S^{private} < S^{FB}$.[29] That is, equilibrium prices in the constrained optimum are not closer to the unconstrained-optimum prices than the decentralized-equilibrium ones, but are instead further away.[30] While this might be surprising at first glance, this result is in fact quite intuitive. Under the first best, the portfolio is observable and it is optimal to choose high-α portfolios. This, of course, will push up the asset prices and reduce expected returns. But, crowded trades are not a problem per se, because a pecuniary externality does not lead to an inefficiency in this case. In contrast, when the contract needs to provide incentives because the portfolio cannot be

[29] The expression for the first-best asset prices is given in Lemma 5 in the Appendix. Comparing it to $S^{private}$ given in Lemma 2 immediately yields the result.

[30] This result parallels that in Lorenzoni (2008), where the decentralized equilibrium falls between the constrained and unconstrained optima in terms of amount of borrowing and asset prices. However, in Lorenzoni's model the inequality signs in the price comparison are reverse—decentralized-equilibrium asset prices are lower than in the constrained optimum (higher in our model) and higher than in the first best (lower in our model).

observed, a pecuniary externality does lead to an inefficiency, and crowded trades pose a problem as they reduce the effectiveness of incentive provision. While the comparison to the first best may not be relevant for practical purposes (as the first best is unattainable), it is helpful to highlight how exactly the mechanism that we explore works.

5 Conclusions

We consider the problem of optimal incentive provision for portfolio managers in a general-equilibrium asset-pricing model. The optimal contacts involve benchmarking. We show that by ignoring the effects of contracts on equilibrium prices, shareholders impose an externality on each other—the effectiveness of their incentive contracts is lower than they perceive it to be. The reason is that contracts incentivize the managers to invest more in stocks with higher α as well as stocks in the benchmark. This boosts prices and lowers returns, making the marginal benefit of α-chasing lower for everyone. The social planner, who internalizes the effects of contracts on equilibrium prices, opts for less incentive provision, less benchmarking, and lower asset-management costs.

As for directions for future research, it could be interesting to incorporate passive asset managers into the model. However, such an extension does present challenges. The existing evidence on the compensation contracts in the asset management industry covers only active funds. Very little is known about contracts of managers in passive funds. Before engaging in modeling of passive managers, it would be important to collect such evidence. A natural starting point would be to analyze the Statements of Additional Information filed by the U.S. mutual funds with the SEC, which contain information on managers' compensation structure. If contracts of passive managers turn out to be incentive contracts, it would be interesting to understand the incentive problem they solve. It is not at all obvious what kind of incentive problem would result in optimal contracts that make the (passive) managers closely follow the benchmark. We leave this problem for future work.

Appendix

Proof of Lemma 1. Equation (5) immediately follows from taking the first-order condition of problem (3) with respect to x. Similarly, (6) follows from taking the first-order condition of problem (4) with respect to x. $\square$

Lemma 5 (First Best). *If x is observable/contractible or if the private cost ψ is zero, then $a = 1/2$ and $b = 0$, and the asset prices are given by $S^{FB} = \mu - \gamma\Sigma\bar{x} + 2\lambda_M(\Delta - \psi)$.*

Proof of Lemma 5. When x is contractible, the problem of the fund investor is simply to maximize $U^F + U^M$. The first-best demand is

$$x^M = \Sigma^{-1}\frac{\Delta - \psi + \mu - S}{\gamma[a^2 + (1-a)^2]} + (2a-1)\frac{b\theta}{a^2 + (1-a)^2}.$$

The first-order condition with respect to $b\theta$ is $\gamma\Sigma(y - z) = 0$. The first-order condition with respect to a is $\gamma\left[\Sigma(y-z)\right]^\top x + \gamma(1-2a)\sigma_\varepsilon^2 = 0$, which, using the first-order condition with respect to $b\theta$ immediately implies $a = 1/2$. Then setting $b = 0$ satisfies the first-order condition with respect to $b\theta$. The first-best demand is then

$$x^M = \Sigma^{-1}\frac{\Delta - \psi + \mu - S}{\gamma/2}.$$

Finally, the expression for equilibrium asset prices in the first best is given by

$$S^{FB} = \mu - \gamma\Sigma\bar{x} + 2\lambda_M(\Delta - \psi).$$

Compared to (22), $S^{FB} > S^{private}$.

Finally, notice that equilibrium holdings of the asset manager are

$$x^M_{FB} = 2\left[\bar{x} + \frac{\Sigma^{-1}}{\gamma}\lambda_D(\Delta - \psi)\right].$$

Notice that if the asset manager holds positive amount of each asset in the first best, then part (a) of Assumption 1 must hold. Therefore part (a) of Assumption 1 is a necessary condition for no short-selling to occur in the first best. $\square$

Proof of Lemma 2. The first-order condition with respect to $b\theta/a$ is

$$\begin{aligned}
0 &= \Delta - \psi + \mu - S - \gamma\Sigma z, \\
0 &= \Delta - \psi + \mu - S - \gamma\Sigma\left[\Sigma^{-1}\frac{\Delta - \psi/a + \mu - S}{\gamma}\left(\frac{1}{a} - 1\right) + \frac{b\theta}{a}\right], \\
\gamma\Sigma\frac{b\theta}{a} &= \Delta - \psi + \mu - S + \left(1 - \frac{1}{a}\right)\left(\Delta - \frac{\psi}{a} + \mu - S\right), \\
\gamma\Sigma\frac{b\theta}{a} &= \left(2 - \frac{1}{a}\right)(\Delta - \psi + \mu - S) + \left(1 - \frac{1}{a}\right)\left(1 - \frac{1}{a}\right)\psi.
\end{aligned}$$

Using the expression for prices given in (7),

$$\gamma\Sigma\frac{b\theta}{a} = \left(2 - \frac{1}{a}\right)\left[\Delta - \psi + \gamma\Sigma\Lambda\left(\bar{x} - \lambda_M\frac{b\theta}{a}\right) - \frac{\lambda_M}{a}\Lambda\left(\Delta - \frac{\psi}{a}\right)\right] + \left(1 - \frac{1}{a}\right)\left(\psi - \frac{\psi}{a}\right).$$

Rearranging terms, and using (8),

$$\begin{aligned}
&\gamma\Sigma\left[1 + \left(2 - \frac{1}{a}\right)\Lambda\lambda_M\right]\frac{b\theta}{a} \\
&\quad = \gamma\Sigma\Lambda\left(2 - \frac{1}{a}\right)\bar{x} + \left(2 - \frac{1}{a}\right)\lambda_D\Lambda(\Delta - \psi) + \left[1 - \frac{1}{a} - \left(2 - \frac{1}{a}\right)\frac{\lambda_M}{a}\Lambda\right]\left(\psi - \frac{\psi}{a}\right), \\
&\gamma\Sigma\Lambda\frac{b\theta}{a} = \Lambda\left(2 - \frac{1}{a}\right)[\gamma\Sigma\bar{x} + \lambda_D(\Delta - \psi)] - \left[\frac{\lambda_M}{a} + \lambda_D\left(\frac{1}{a} - 1\right)\right]\Lambda\left(\psi - \frac{\psi}{a}\right), \qquad (34) \\
&b\theta = (2a - 1)\left[\bar{x} + \frac{\Sigma^{-1}}{\gamma}\lambda_D(\Delta - \psi)\right] + (1 - a)\left[\frac{1}{a} - \left(\frac{\lambda_M}{a} + \lambda_D\right)\right]\frac{\Sigma^{-1}}{\gamma}\psi.
\end{aligned}$$

Similarly, the first-order condition for b and θ evaluated at the equilibrium is given by (20) and (21), respectively.

Plugging (34) into the expression for prices (7), obtain

$$\begin{aligned}
S^{private} &= \mu - \gamma\Sigma\Lambda\bar{x} + \lambda_M\gamma\Sigma\Lambda b\theta + \Lambda\frac{\lambda_M}{a}\left(\Delta - \frac{\psi}{a}\right) \\
&= \mu - \gamma\Sigma\Lambda\bar{x}\left(1 - \lambda_M\left(2 - \frac{1}{a}\right)\right) + \lambda_M(\Delta - \psi) + \frac{\lambda_M\Lambda}{a}[\lambda_M + a\lambda_D]\left(\Delta - \frac{\psi}{a}\right) \\
&= \mu - \gamma\Sigma\bar{x} + \lambda_M\left(\Delta - \psi + \Delta - \frac{\psi}{a}\right).
\end{aligned}$$

Substituting (16) into (6) and rearranging terms, obtain

$$\gamma\Sigma x^M = (\Delta - \psi + \mu - S) + \left(\Delta - \frac{\psi}{a} + \mu - S\right).$$

Substituting the expression for prices $S = S^{private}$ derived above and rearranging terms yields (23). $\square$

Proof of Proposition 1. The result immediately follows from (20) and Assumption 1. $\square$

Proof of Lemma 3. Denote the (k,ℓ)-th element of matrix Σ^{-1} by $e_{k,\ell}$, where $e_{k,\ell} = e_{\ell,k}$ by symmetry, in particular, $e_{i,j} = e_{j,i}$. Since assets i and j are assumed to be identical (except for Δ's and ψ's), we have $e_{i,i} = e_{j,j}$ and $e_{i,k} = e_{j,k}$ for all $k \neq i, j$ (i.e., assets i and j have the same variance and covariance with other assets). As a result,

$$\theta_i - \theta_j = (e_{i,i} - e_{i,j})\left\{\frac{2a-1}{b\gamma}\lambda_D[\Delta_i - \psi_i - (\Delta_j - \psi_j)] + \frac{1-a}{b\gamma}\left(\frac{1}{a} - \frac{\lambda_M}{a} - \lambda_D\right)(\psi_i - \psi_j)\right\}.$$

Because Σ^{-1} is positive definite, we have $e_{i,i} > 0,\ e_{i,i}e_{j,j} - e_{i,j}^2 > 0,\ e_{i,i} > |e_{i,j}|$. As a result, $e_{i,i} - e_{i,j} > 0$, and thus $\theta_i > \theta_j$ whenever $\Delta_i - \psi_i \geq \Delta_j - \psi_j$, $\psi_i \geq \psi_j$, and at least one of the inequalities is strict. With a slight modification, this proof also applies to the socially-optimal contract. $\square$

The Social Planner's Problem in Terms of Exponential Utilities:

$$\begin{aligned}\max_{a,b,\theta,c,x=x^M,x^D} & -\tilde{\omega}_F E\exp\left\{-\gamma\left[\left(x_{-1}^F\right)^\top S + r_x - (ar_x - br_{\mathbf{b}}) - c\right]\right\} \\ & -\tilde{\omega}_D E\exp\left\{-\gamma\left[\left(x_{-1}^D\right)^\top S + \left(x^D\right)^\top (D - S)\right]\right\}\end{aligned}$$

subject to (5), (6), and (9), where $\tilde{\omega}_i$, $i = S, C$, are the modified Pareto weights. From the first-order condition with respect to c it follows that the Lagrange multiplier on the participation constraint equals $\tilde{\omega}_F MU_F/MU_M$, where MU_i denotes the expected marginal utility of agent i. This value is the effective Pareto weight on the asset manager's utility given that transfers between the fund investor and manager are allowed. Similarly, if transfers between fund investors and direct investors were allowed, then $\tilde{\omega}_F MU_F = \tilde{\omega}_D MU_D$, and the redistribution effects would be zero. And without transfers, the Pareto weights that cancel out the redistribution effects (in the formulation with exponential utilities) are equal to inverse marginal utilities, $\tilde{\omega}_i = 1/MU_i$.

Proof of Lemma 4.

The social planner's first-order condition with respect to $b\theta/a$ is[31]

$$\left[\omega_F \left(x^F_{-1} - x^M\right)^\top + \omega_D \left(x^D_{-1} - x^D\right)^\top\right] \frac{\partial S}{\partial (b\theta/a)}$$
$$+ \omega_F \left[(\Delta - \psi + \mu - S - \gamma\Sigma z)^\top + (\Delta + \mu - S - \gamma\Sigma z)^\top \left(\frac{1}{a} - 1\right) \frac{\partial y}{\partial S} \frac{\partial S}{\partial (b\theta/a)}\right] = 0.$$

Canceling out the redistribution effects and using $\partial y/\partial S = -\Sigma^{-1}/\gamma$ and $\partial S/\partial(b\theta/a) = \gamma\Sigma\Lambda\lambda_M$, the above equation (or (25)) becomes

$$\begin{aligned}
0 &= \Delta - \psi + \mu - S - \gamma\Sigma z - (\Delta + \mu - S - \gamma\Sigma z) \frac{1-a}{a} \Lambda\lambda_M, \\
0 &= \Delta - \psi + \mu - S - \gamma\Sigma z - \psi \frac{(1/a - 1)\Lambda\lambda_M}{1 - (1/a-1)\Lambda\lambda_M}, \\
0 &= \Delta - \psi + \mu - S - \gamma\Sigma \left[\Sigma^{-1} \frac{\Delta - \psi/a + \mu - S}{\gamma} \left(\frac{1}{a} - 1\right) + \frac{b\theta}{a}\right] - \psi \frac{(1/a - 1)\Lambda\lambda_M}{1 - (1/a-1)\Lambda\lambda_M}.
\end{aligned}$$

Rearranging terms,

$$\begin{aligned}
\gamma\Sigma\frac{b\theta}{a} &= \Delta - \psi + \mu - S + \left(1 - \frac{1}{a}\right)\left(\Delta - \frac{\psi}{a} + \mu - S\right) - \psi \frac{(1-a)/a\Lambda\lambda_M}{1 - (1/a-1)\Lambda\lambda_M}, \\
\gamma\Sigma b\theta &= (2a-1)(\Delta - \psi + \mu - S) + (1-a)\left[\frac{1-a}{a} - \frac{\Lambda\lambda_M}{1 - (1/a-1)\Lambda\lambda_M}\right]\psi, \\
\gamma\Sigma b\theta &= (2a-1)(\Delta - \psi + \mu - S) + (1-a)\left(\frac{1}{a} - \frac{1}{\lambda_M + \lambda_D}\right)\psi. \qquad (35)
\end{aligned}$$

Alternatively,

$$\begin{aligned}
0 &= \Delta - \frac{\lambda_M/a + \lambda_D}{\lambda_M + \lambda_D}\psi + \mu - S - \gamma\Sigma z, \\
\gamma\Sigma\frac{b\theta}{a} &= \Delta - \frac{\lambda_M/a + \lambda_D}{\lambda_M + \lambda_D}\psi + \mu - S + \left(1 - \frac{1}{a}\right)\left(\Delta - \frac{\psi}{a} + \mu - S\right), \\
\gamma\Sigma b\theta &= (2a-1)\left(\Delta - \frac{\lambda_M/a + \lambda_D}{\lambda_M + \lambda_D}\psi + \mu - S\right) + (1-a)\left(\frac{1}{a} - \frac{\lambda_M/a + \lambda_D}{\lambda_M + \lambda_D}\right)\psi.
\end{aligned}$$

[31] Because the asset manager's and direct investor's utilities are maximized with respect to y and x^D, respectively, by the Envelope theorem the only terms from their payoffs that enter the first-order conditions are those entering the redistribution term.

Substituting the expression for prices into (35) leads to

$$b\theta = (2a-1)\left[\bar{x} + \frac{\Sigma^{-1}}{\gamma}\lambda_D(\Delta - \psi)\right] + (1-a)\left[\frac{1}{a} - \frac{\lambda_M/a + \lambda_D}{\lambda_M + \lambda_D}\right]\frac{\Sigma^{-1}}{\gamma}\psi,$$

$$b\theta = (2a-1)\left[\bar{x} + \frac{\Sigma^{-1}}{\gamma}\lambda_D(\Delta - \psi)\right] + (1-a)\left[\frac{1}{a} - 1\right]\frac{\lambda_D}{\lambda_M + \lambda_D}\frac{\Sigma^{-1}}{\gamma}\psi.$$

Similarly, the first-order conditions for b and θ separately, evaluated in equilibrium, are then given by (30) and (31).

Following the same steps as in the proof of Lemma 2, the expression for prices is

$$S^{social} = \mu - \gamma\Sigma\bar{x} + \lambda_M(\Delta - \psi) + \lambda_M\left(\Delta - \frac{\psi}{a}\right) - \left(\frac{1}{a} - 1\right)\frac{\lambda_M^2}{\lambda_M + \lambda_D}\psi,$$

or, alternatively,

$$S^{social} = \mu - \gamma\Sigma\bar{x} + \lambda_M\left(\Delta - \frac{\lambda_M/a + \lambda_D}{\lambda_M + \lambda_D}\psi + \Delta - \frac{\psi}{a}\right).$$

Finally, substituting (27) into (6) and rearranging terms, obtain

$$\gamma\Sigma x^M = \left[\Delta - \frac{\lambda_M/a + \lambda_D}{\lambda_M + \lambda_D}\psi + \mu - S\right] + \left[\Delta - \frac{\psi}{a} + \mu - S\right].$$

Substituting (32) and rearranging terms yields (33). $\square$

Proof of Proposition 2.

(a) Comparison $a^{social} < a^{private}$ follows straightforwardly from comparing (19) and (29).

(b) Denote $a_1 = a^{private}$ and $a_2 = a^{social}$. From the first-order conditions with respect to b evaluated in equilibrium,

$$b_1 = (2a_1 - 1)\mathbf{1}^\top\left[\bar{x} + \frac{\Sigma^{-1}}{\gamma}\lambda_D(\Delta - \psi)\right] + (1-a_1)\left[\frac{1}{a_1} - \frac{\lambda_M}{a_1} - \lambda_D\right]\mathbf{1}^\top\frac{\Sigma^{-1}}{\gamma}\psi,$$

$$b_2 = (2a_2 - 1)\mathbf{1}^\top\left[\bar{x} + \frac{\Sigma^{-1}}{\gamma}\lambda_D(\Delta - \psi)\right]$$
$$+ (1-a_2)\left[\frac{1}{a_2} - \frac{1}{\lambda_M + \lambda_D}\left(\frac{\lambda_M}{a_2} + \lambda_D\right)\right]\mathbf{1}^\top\frac{\Sigma^{-1}}{\gamma}\psi.$$

Under Assumption 1, in order to show that $b_1 > b_2$, it is sufficient to show that

$$(1-a_1)\left[\frac{1}{a_1}-\frac{\lambda_M}{a_1}-\lambda_D\right] > (1-a_2)\left[\frac{1}{a_2}-\frac{1}{\lambda_M+\lambda_D}\left(\frac{\lambda_M}{a_2}+\lambda_D\right)\right],$$

which (given that both sides of the above inequality are positive) is equivalent to

$$\frac{(1-a_1)/a_1}{(1-a_2)/a_2}\frac{\lambda_M+a_1\lambda_D+(1-2a_1)\lambda_D}{(\lambda_M+a_2\lambda_D)\lambda_D/(\lambda_M+\lambda_D)+(1-2a_2)\lambda_D} > 1. \tag{36}$$

Given the equilibrium conditions for a_1 and a_2, we have

$$\frac{1-a_1}{a_1^3(2a_1-1)} = \frac{1-a_2}{a_2^3(2a_2-1)}\frac{\lambda_D}{\lambda_M+\lambda_D}. \tag{37}$$

Substitute this into inequality (36), obtain

$$\frac{a_1^2(2a_1-1)}{a_2^2(2a_2-1)}\frac{\lambda_D}{\lambda_M+\lambda_D}\frac{\lambda_M+a_1\lambda_D+(1-2a_1)\lambda_D}{(\lambda_M+a_2\lambda_D)\lambda_D/(\lambda_M+\lambda_D)+(1-2a_2)\lambda_D} > 1.$$

Since $a_1 > a_2$, it suffices to show that

$$\begin{aligned}
&\frac{\lambda_M+a_1\lambda_D+(1-2a_1)\lambda_D}{(\lambda_M+a_2\lambda_D)\lambda_D/(\lambda_M+\lambda_D)+(1-2a_2)\lambda_D} > \frac{\lambda_D+\lambda_M}{\lambda_D}\\
\Leftrightarrow\ &\lambda_M+a_1\lambda_D+(1-2a_1)\lambda_D > \lambda_M+a_2\lambda_D+(1-2a_2)(\lambda_D+\lambda_M)\\
\Leftrightarrow\ &\lambda_M(2a_2-1) > \lambda_D(a_1-a_2). \qquad (38)
\end{aligned}$$

To show (38), use equation (37), rearranging which we get

$$\frac{1-a_1}{a_1^3(2a_1-1)}\frac{\lambda_M}{\lambda_D} = \frac{1-a_2}{a_2^3(2a_2-1)} - \frac{1-a_1}{a_1^3(2a_1-1)},$$

or, equivalently,

$$\frac{\lambda_M(2a_2-1)}{\lambda_D} = \frac{a_1^3}{1-a_1}\left[\frac{(1-a_2)(2a_1-1)}{a_2^3} - \frac{(1-a_1)(2a_2-1)}{a_1^3}\right].$$

The right-hand side of the above equation equals

$$
\begin{aligned}
&\frac{-a_1^3+2a_1^4-2a_1^4a_2+a_2a_1^3-(-a_2^3+2a_2^4-2a_2^4a_1+a_1a_2^3)}{(1-a_1)a_2^3}\\
&=\frac{(a_1-a_2)}{(1-a_1)a_2^3}[-(1+2a_1a_2)(a_1^2+a_1a_2+a_2^2)+2(a_1+a_2)(a_1^2+a_2^2)+a_1a_2(a_1+a_2)].
\end{aligned}
$$

Rearranging terms and doing some more algebra, yields

$$
\begin{aligned}
&\frac{\lambda_M(2a_2-1)}{\lambda_D(a_1-a_2)}\\
&=\frac{(2a_1-1)a_1^2(1-a_2)+(2a_2-1)a_2^2(1-a_1)+(2a_1-1)a_1a_2+2a_1a_2^2(1-a_1)}{a_2^3(1-a_1)}.
\end{aligned}
$$

Since $1/2<a_2<a_1<1$,

$$
\begin{aligned}
&\frac{\lambda_M(2a_2-1)}{\lambda_D(a_1-a_2)}\\
&>\frac{(2a_1-1)a_1^2(1-a_2)+(2a_2-1)a_2^2(1-a_1)+(2a_1-1)a_1a_2+a_2^3(1-a_1)}{a_2^3(1-a_1)}>1.
\end{aligned}
$$

Thus $b_1>b_2$. □

Proof of Proposition 3.

(a) Follows immediately from comparing (22) and (32) and using part (a) of Proposition 2:

$$S^{private}-S^{social}=\lambda_M\left(\frac{1}{a^{social}}-\frac{1}{a^{private}}\right)\psi+\left(\frac{1}{a^{social}}-1\right)\frac{\lambda_M^2}{\lambda_M+\lambda_D}\psi,$$

both terms of which are strictly positive.

(b) Using (23) and (33),

$$\psi^\top\left(x^M_{social}-x^M_{private}\right)=\lambda_D\psi^\top\frac{\Sigma^{-1}}{\gamma}\psi\left[1-\frac{\lambda_M/a^{social}+\lambda_D}{\lambda_M+\lambda_D}+\frac{1}{a^{social}}-\frac{1}{a^{private}}\right].$$

Since Σ^{-1} is positive definite and the expression in square brackets is negative (as $a^{social}<a^{private}<1$), we have $\psi^\top\left(x^M_{social}-x^M_{private}\right)<0$. □

Proof of Claim 1. Denote $a_1 = a^{private}$ and $a_2 = a^{social}$, and let $e_{i,j}$ be the (i,j)-th element of matrix Σ^{-1} as defined in the proof of Lemma 3. Then

$$\begin{aligned}
\theta_i^{private} - \theta_j^{private} &= \frac{2a_1 - 1}{b_1 \gamma} \lambda_D (e_{i,i} - e_{i,j})[\Delta_i - \Delta_j - (\psi_i - \psi_j)] \\
&\quad + \frac{1 - a_1}{b_1 \gamma} \left(\frac{1}{a_1} - \frac{\lambda_M}{a_1} - \lambda_D \right) (e_{i,i} - e_{i,j})(\psi_i - \psi_j), \\
\theta_i^{social} - \theta_j^{social} &= \frac{2a_2 - 1}{b_2 \gamma} \lambda_D (e_{i,i} - e_{i,j})[\Delta_i - \Delta_j - (\psi_i - \psi_j)] \\
&\quad + \frac{1 - a_2}{b_2 \gamma} \left(\frac{1}{a_2} - \frac{\lambda_M / a_2 + \lambda_D}{\lambda_M + \lambda_D} \right) (e_{i,i} - e_{i,j})(\psi_i - \psi_j).
\end{aligned}$$

Using similar steps as in the proof of $b_1 > b_2$ in part (b) of Proposition 2 we can show that $b_1/(2a_1 - 1) > b_2/(2a_2 - 1)$ and thus $(2a_1 - 1)/b_1 < (2a_2 - 1)/b_2$. Furthermore,

$$\begin{aligned}
&\frac{1 - a_1}{b_1} \left[\frac{1}{a_1} - \frac{\lambda_M}{a_1} - \lambda_D \right] \\
&= \left[\mathbf{1}^\top \frac{\Sigma^{-1}}{\gamma} \psi + \frac{2a_1 - 1}{(1 - a_1)(1/a_1 - \lambda_M/a_1 - \lambda_D)} \mathbf{1}^\top \left(\bar{x} + \frac{\Sigma^{-1}}{\gamma} \lambda_D (\Delta - \psi) \right) \right]^{-1}, \\
&\frac{1 - a_2}{b_2} \left[\frac{1}{a_2} - \frac{\lambda_M / a_2 + \lambda_D}{\lambda_M + \lambda_D} \right] \\
&= \left[\mathbf{1}^\top \frac{\Sigma^{-1}}{\gamma} \psi + \frac{2a_2 - 1}{(1 - a_2)\left[1/a_2 - (\lambda_M/a_2 + \lambda_D)/(\lambda_M + \lambda_D)\right]} \mathbf{1}^\top \left(\bar{x} + \frac{\Sigma^{-1}}{\gamma} \lambda_D (\Delta - \psi) \right) \right]^{-1}.
\end{aligned}$$

From the proof of $b_1 > b_2$ in part (b) of Proposition 2 we know that

$$\frac{2a_1 - 1}{(1 - a_1)(1/a_1 - \lambda_M/a_1 - \lambda_D)} < \frac{2a_2 - 1}{(1 - a_2)\left[1/a_2 - (\lambda_M/a_2 + \lambda_D)/(\lambda_M + \lambda_D)\right]},$$

and thus

$$\frac{1 - a_1}{b_1} \left[\frac{1}{a_1} - \frac{\lambda_M}{a_1} - \lambda_D \right] > \frac{1 - a_2}{b_2} \left[\frac{1}{a_2} - \frac{\lambda_M / a_2 + \lambda_D}{\lambda_M + \lambda_D} \right].$$

Hence when $\Delta_i - \psi_i \geq \Delta_j - \psi_j$ and $\psi_i \leq \psi_j$, and at least one inequality is strict, we have $\theta_i^{social} - \theta_j^{social} > \theta_i^{private} - \theta_j^{private}$. And reverse, if $\Delta_i - \psi_i \leq \Delta_j - \psi_j$ and $\psi_i \geq \psi_j$, and at least one inequality is strict, then we have $\theta_i^{social} - \theta_j^{social} < \theta_i^{private} - \theta_j^{private}$. The interpretation is that the socially-optimal contract puts relatively less weight on incentive provision and thus relatively more weight on protecting the asset manager from risk. $\square$

References

Acemoglu, D., and A. Simsek (2012): "Moral Hazard and Efficiency in General Equilibrium with Anonymous Trading," MIT working paper.

Adams, J. C., S. A. Mansi, and T. Nishikawa (2014): "Affiliated Agents, Boards of Directors and Mutual Fund Securities Lending Returns," *The Journal of Financial Research*, 37(4), 461–493.

Admati, A. R., and P. Pfleiderer (1997): "Does It All Add Up? Benchmarks and the Compensation of Active Portfolio Managers," *Journal of Business*, 70(3), 323–350.

Albuquerque, R. A., L. M. B. Cabral, and J. Guedes (forthcoming): "Incentive Pay and Systemic Risk," *Review of Financial Studies*.

Anand, A., C. Jotikasthira, and K. Venkataraman (2018): "Do Buy-side Institutions Supply Liquidity in Bond Markets? Evidence from Mutual Funds," *SMU Cox School of Business Research Paper*.

Bank for International Settlements (2003): "Incentive Structures in Institutional Asset Management and Their Implications for Financial Markets," March 31, 2013, https://www.bis.org/publ/cgfs21.htm.

Basak, S., and A. Pavlova (2013): "Asset Prices and Institutional Investors," *American Economic Review*, 103(5), 1728–1758.

Berk, J. B., and R. C. Green (2004): "Mutual Fund Flows and Performance in Rational Markets," *Journal of Political Economy*, 112(6), 1269–1295.

Biais, B., F. Heider, and M. Hoerova (2019): "Variation Margins, Fire Sales, and Information-Constrained Optimality," ECB Working Paper.

Brennan, M. (1993): "Agency and Asset Pricing," UCLA Working Paper.

Buffa, A., D. Vayanos, and P. Woolley (2014): "Asset Management Contracts and Equilibrium Prices," NBER Working Paper 20480.

Cuoco, D., and R. Kaniel (2011): "Equilibrium Prices in the Presence of Delegated Portfolio Management," *Journal of Financial Economics*, 101(2), 264–296.

Cvitanic, J., and H. Xing (2018): “Asset Pricing Under Optimal Contracts,” *Journal of Economic Theory*, 173(C), 142–180.

Davila, E., and A. Korinek (2018): “Pecuniary Externalities in Economies with Financial Frictions,” *Review of Economic Studies*, 85, 352–395.

Eisele, A., T. Nefedova, G. Parise, and K. Peijnenburg (forthcoming): “Trading Out of Sight: An Analysis of Cross-Trading in Mutual Fund Families,” Journal of Financial Economics.

Fama, E. F., and K. R. French (2010): “Luck versus Skill in the Cross-Section of Mutual Fund Returns,” *The Journal of Finance*, 65(5), 1915–1947.

Fershtman, C., and K. L. Judd (1987): “Equilibrium Incentives in Oligopoly,” *The American Economic Review*, 77, 927–940.

Geanakoplos, J., and H. M. Polemarchakis (1996): “Existence, Regularity, and Constrained Suboptimality of Competitive Allocations When the Asset Market Is Incomplete,” in Gerard Debreu, ed., *General Equilibrium Theory, Vol. 2*, Cheltenham, UK: Elgar.

Greenwald, B. C., and J. E. Stiglitz (1986): “Externalities in Economies with Imperfect Information and Incomplete Markets,” *Quarterly Journal of Economics*, 90, 229?264.

Gromb, D., and D. Vayanos (2002): “Equilibrium and Welfare in Markets with Financially Constrained Arbitrageurs,” *Journal of Financial Economics*, 66(2-3), 361–407.

Hart, O. D. (1975): “On the Optimality of Equilibrium When the Market Structure is Incomplete,” *Journal of Economic Theory*, 11, 418–443.

He, Z., and P. Kondor (2016): “Inefficient Investment Waves,” *Econometrica*, 84(2), 735–780.

Holmstrom, B. (1979): “Moral Hazard and Observability,” *The Bell Journal of Economics*, 10(1), 74–91.

Holmstrom, B., and P. Milgrom (1987): “Aggregation and Linearity in the Provision of Intertemporal Incentives,” *Econometrica*, 55(2), 303–328.

——— (1991): "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design," *Journal of Law, Economics, and Organization*, 7(Special Issue), 24–52.

Ibert, M., R. Kaniel, S. Van Nieuwerburgh, and R. Vestman (2017): "Are Mutual Fund Managers Paid for Investment Skill?," *The Review of Financial Studies*, 31(2), 715–772.

Johnson, T. L., and G. Weitzner (2019): "Distortions Caused by Lending Fee Retention," University of Texas at Austin Working Paper.

Kacperczyk, M., C. Sialm, and L. Zheng (2008): "Unobserved Actions of Mutual Funds," *Review of Financial Studies*, 21(6), 2379–2416.

Kashyap, A., N. Kovrijnykh, J. Li, and A. Pavlova (2018): "The Benchmark Inclusion Subsidy," NBER Working Paper 25337.

Keim, D. B. (1999): "An Analysis of Mutual Fund Design: the Case of Investing in Small-Cap Stocks," *Journal of Financial Economics*, 51, 173–194.

Khorana, A., H. Servaes, and L. Wedge (2007): "Portfolio Manager Ownership and Fund Performance," *Journal of Financial Economics*, 85, 179–204.

Lorenzoni, G. (2008): "Inefficient Credit Booms," *The Review of Economic Studies*, 75.

Ma, L., Y. Tang, and J.-P. Gómez (2019): "Portfolio Manager Compensation in the U.S. Mutual Fund Industry," *Journal of Finance*, 70, 587–638.

Ozdenoren, E., and K. Yuan (2017): "Contractual Externalities and Systemic Risk," *Review of Economic Studies*, 84, 1789–1817.

Rinne, K., and M. Suominen (2016): "Short-term Reversals, Returns to Liquidity Provision and the Costs of Immediacy," *Working paper*.

Stoughton, N. M. (1993): "Moral Hazard and the Portfolio Management Problem," *Journal of Finance*, 48(5), 2009–2028.

Sung, J. (1995): "Linearity with Project Selection and Controllable Diffusion Rate in Continuous-Time Principal-Agent Problems," *The RAND Journal of Economics*, 26(4), 720–743.