Delegated Blocks*

Amil Dasgupta† Richmond Mathews‡

This version: February 2025

Abstract

Will asset managers with large amounts of capital and high risk-bearing capacity hold large blocks and monitor aggressively? Both block size and monitoring intensity are governed by the contractual incentives of asset managers, which are themselves endogenous. We show that when high risk-bearing capacity arises via optimal delegation, funds hold smaller blocks and monitor significantly less than proprietary investors with identical risk-bearing capacity. This is because the optimal contract enables the separation of risk sharing and monitoring incentives. Our findings rationalize characteristics of real-world asset managers and imply that block sizes will be a poor predictor of monitoring intensity.

*We are grateful to Anat Admati, Patrick Bolton, Francesco Celentano, Pierre Chaigneau, Alvin Chen, Joanne Chen, Kenneth Deng, Yunzhi Hu, Mark Loewenstein, Nadya Malenko, John Moore, Shri Santosh, Dimitri Vayanos, Vladimir Vladimirov, Josef Zechner and conference and seminar audiences at CUHK, HEC Paris, HKU, HKUST, LSE, Maryland, Nanyang Business School, NUS, SMU, the Cambrige Corporate Finance Theory Symposium (2024), EFA (2023), FIRS (2023), IBEO conference (2023), London FIT conference (2024), OxFIT (2024), SHoF Corporate Governance Conference (2024), Vienna Festival of Finance Theory (2024), and WFA (2023) for helpful comments. Dasgupta acknowledges financial support from the ESRC via Research Grant ES/S016686/1.

†London School of Economics and ECGI. Email: a.dasgupta@lse.ac.uk

‡Smith School of Business, University of Maryland at College Park. Email: rmathews@umd.edu

1 Introduction

The rise of asset managers has led to the concentration of vast amounts of capital in the hands of institutional investors.[1] How is this likely to affect block ownership and corporate governance? These investors have large risk-bearing capacity and thus are able—in principle—to hold large blocks and monitor portfolio firms aggressively. However, both block size and the extent of monitoring are endogenous to the contractual incentives of asset managers. Such contractual incentives—in turn—are endogenously determined, and will anticipate ownership and monitoring decisions. Will institutional investors hold large blocks commensurate to their risk-bearing capacity in equilibrium? Conditional on holding such blocks, will they monitor firms aggressively? To help answer these questions, we study the economics of delegated blockholding. In particular, we characterize blockholder monitoring and risk sharing in markets where equity ownership is optimally delegated, and both equity block sizes and the level of monitoring are determined by endogenous contracts established between asset managers and their clients.

The existing literature focuses on the formation of long-term *proprietary* blocks. Admati, Pfleiderer, and Zechner (1994)—APZ henceforth—take as given the existence of a proprietary trader with high risk-bearing capacity and consider whether anticipated monitoring costs will limit the trader's willingness to hold large blocks. Under broad and plausible conditions, they find the answer is "no"—as long as traders with high risk-bearing capacity cannot commit to limit their trading, they will trade to the competitive risk sharing allocation and monitor at a level consistent with that allocation. This is because the ability to trade repeatedly erodes the large trader's strategic advantage. APZ's striking finding is confirmed in the fully dynamic analysis of DeMarzo and Urosevic (2006). Overall, therefore, the existing literature provides a reassuring view: anticipated monitoring costs will not limit large traders' willingness to hold large blocks and monitor.

We show that when high risk-bearing capacity is instead attained endogenously via delegation—wherein agents without full access to financial markets hire professional asset

[1] See, e.g., Dasgupta, Fos, and Sautner (2021) for relevant stylized facts.

managers to trade and monitor for them—outcomes are dramatically different. First, the optimal fund holds less of the risky asset, i.e., a smaller block, than an investor with the same risk-bearing capacity would under the competitive risk sharing allocation. Second, delegation separates block sizes and monitoring incentives, because—within a fund—monitoring is undertaken by professional asset managers. It is *their* stake, not the overall block size, that determines the fund's level of monitoring. The optimal delegation contract allocates a stake to these professional asset managers that results in a level of monitoring that would be privately optimal for fund investors at their *initial* endowment. These two effects combined imply that the optimal fund undertakes significantly less monitoring than a proprietary blockholder of identical risk-bearing capacity. Overall, delegated blockholding delivers less monitoring and inferior risk sharing relative to the proprietary blocks benchmark, but gives rise to monitoring and risk sharing benefits when equity blocks would not otherwise exist.

Model summary. To establish the proprietary blocks benchmark, we start with a version of APZ's classical "CARA-Normal" model featuring a firm with Normally distributed equity cash flows and a group of investors with CARA utility. There are two types of investors: a single large entity, L, with risk tolerance, i.e., risk-bearing capacity, of λ, and a continuum of small investors with aggregate risk-bearing capacity of $1 - \lambda$.[2] In addition to trading (potentially many times) in a Walrasian market, L can also monitor the firm. Such monitoring is costly for L but increases cash flows to all equity holders. The competitive equilibrium allocation in such an economy would involve L holding λ fraction of the firm's equity.

Imagine that $L's$ initial endowment of the risky asset is $\omega < \lambda$. Will L trade from ω all the way to λ? There are several impediments. First, L knows that if she trades to λ she will then monitor at a commensurately higher intensity and all $1 - \lambda$ other shareholders will benefit from such monitoring. Second, L knows that along the way to λ she must pay the full value of future monitoring when acquiring shares, i.e., she moves prices against herself as she trades. However, in a key result, APZ show that as long as L cannot commit to limit her trading,

[2] For expositional ease, in the introduction we describe our model "as if" the economy has unit aggregate risk-bearing capacity. However, our formal analysis is valid for any arbitrary aggregate risk-bearing capacity.

she will nevertheless trade all the way to λ and monitor at the intensity corresponding to such holdings. This arises because of an endowment effect. Counterfactually, if any sequence of trades led to a final holding level for L that is strictly below λ, she would be tempted to buy a bit more because the current holding is now part of her endowment. Starting from this endowment, there will always be at least some incremental risk sharing gains by buying a bit more, despite having to pay the full value of future monitoring in making such purchases. This result implies that the anticipation of future monitoring costs will *not* preclude investors with high risk-bearing capacity from holding large positions and monitoring accordingly.

We next enrich the model to incorporate delegated blockholding. To model blocks held by funds, we assume that a measure $1-\lambda$ of agents are "financial" investors while the remaining λ measure of agents are "nonfinancial" investors. Financial investors are professionals who specialize in trading and monitoring on their own behalf and, potentially, on behalf of nonfinancial investors via "funds." They are endowed with $1-\omega$ shares, distributed equally among themselves. Nonfinancial investors are otherwise occupied and face a high opportunity cost of participating directly in financial markets to trade or monitor. As a group, they are endowed with $\omega < \lambda$ shares, distributed equally among themselves. Nonfinancial investors thus have a lower endowment of the risky asset than their desired holdings in a competitive equilibrium, and—given the opportunity to participate in financial markets by joining a fund—could enjoy risk sharing benefits by increasing their risky asset exposure.

To model fund formation, we introduce a "fund designer" who offers a contract to all nonfinancial investors—who become Fund Clients (FCs) if they agree to join—and a measure of financial investors—who become Fund Managers (FMs) if they agree to join. Investors join the fund and coinvest only if they find it incentive compatible. The FMs then make trading and monitoring decisions on behalf of the fund based on their contractual incentives.[3] As in the APZ benchmark, the FMs cannot commit to a given trading strategy or monitoring level up front—they always behave opportunistically once the fund has been established.

These assumptions capture key aspects of real world financial markets. In such markets

[3]By assuming that FMs choose their trading strategies, we exclude pure index funds from our formal model. However, as discussed later in the paper, our results have some indirect implications for the role of index funds in governance.

there are professional asset managers, investors who trade only via those professional managers, and investors who trade directly on their own account. In our model, professional asset managers are represented by FMs, investors who trade only via asset managers are represented by FCs, and those who trade directly are represented by financial investors who do not become FMs.

Main results. We have a principal-agent model in which the FCs are the principals and the FMs are agents. For the majority of our analysis, we take the conventional approach in that we solve for the optimal fund as the contract that maximizes the payoffs of the principal, i.e., the FCs. We derive an optimal contract from the point of view of the FCs in two steps. First, we find the full-commitment, privately optimal allocation from the FCs' point of view *as if* they could trade and monitor *and also* commit to specific monitoring and trading strategies. Next, we show that under intuitive conditions a simple fund contract exists that fully achieves this outcome for the FCs, despite the fact that—exactly as in APZ—the FMs cannot commit to specific trading or monitoring strategies. The optimal contract specifies a measure τ of FMs, a fee f paid by FCs to join the fund, and a "skin in the game" parameter $\phi \in [0, 1]$ representing the FMs' share of the fund's assets. Since FMs can choose to unilaterally deviate from the fund and benefit from its monitoring efforts, the contractual fees must compensate FMs for their monitoring costs. Similarly, since FCs can choose to unilaterally deviate from the fund, avoid paying fees, and gain from its monitoring efforts, investing in the fund must deliver sufficient risk sharing benefits to justify these fees. Subject to these two incentive compatibility conditions, the contract aims to induce FMs to trade and monitor so as to achieve the outcome desired by the FCs.

We show that the optimal contract induces radically different trading and monitoring choices relative to the APZ benchmark. A key insight is that delegation separates monitoring incentives from overall holdings. This is because delegated monitoring is undertaken by professional asset managers on behalf of the fund: It is *their* stake, not the fund's overall holdings, that determines the fund's level of monitoring. The optimal contract allocates a share of the fund's assets to FMs that induces monitoring at a level consistent with only the FCs' *initial* endowment; in other words, FCs do not have to compensate FMs for any monitoring that is excessive from the FCs' private perspective. However, since FCs' initial endowment is $\omega < \lambda$, whereas the aggregate risk-bearing capacity of the FC's is λ, the optimal fund monitors less than a proprietary trader with identical risk-bearing capacity in the APZ benchmark. Further, we show that the fund also holds too small an overall position in the asset: in particular, under the optimal contract, the FCs hold a position within the fund that fully reflects their market power as a strategic trader with aggregate risk-bearing capacity λ. Overall, therefore, by separating monitoring incentives from risk sharing, the optimal contract has the flavor of a commitment device that enables FCs to attain their *privately* optimal, full-commitment, levels of both monitoring and risk sharing. But this is attained at the expense of lower overall levels of monitoring and risk sharing in the market relative to a benchmark with large proprietary traders. However, access to financial markets via delegation does enhance both risk sharing and monitoring when equity blocks would not otherwise exist.

No commitment to contracts. While the optimal contract is somewhat reminiscent of a commitment device, we do *not* require long-term commitment to the fund contract. In section 4.1, we analyze a situation where the optimal contract is signed and the fund trades, but then an unexpected opportunity arises ex post to dissolve the existing fund and start a new one that would trade further and monitor at a level closer to the APZ benchmark. We show that no sequence of such ex post recontracting opportunities can revive the APZ result. This is because, unlike retrading (which is always feasible), recontracting is limited by free-riding: once the fund's stake gets large enough, FC's individually get "too close" to their risk sharing optimum and would prefer to leave the fund and free-ride on its monitoring.

Extensions. We extend our analysis in several ways and show that even worse governance outcomes can arise in realistic situations. First, we consider contracts that are optimal from the perspective of agents (i.e., financial investors) by analyzing funds that maximize the rents that can be extracted from nonfinancial investors, subject to their willingness to join.

We show that such funds undertake no monitoring at all and give FCs the same degree of risk sharing as our benchmark optimal fund, thus resulting in a worsening of corporate governance. Second, we allow for perfect competition among groups of FMs. In this case, we show that delegation can—again—lead to even worse outcomes from a corporate governance perspective than our optimal fund. However, we also show that our optimal fund can still arise in the presence of such competition given some realistic additional frictions.

Applied implications. Our main results characterize the economics of monitoring and risk sharing in financial markets with delegated blockholding. Given the preponderance of delegated asset managers in modern financial markets, these results are relevant to interpreting key features of blockholding and monitoring that are prevalent today. Specifically, our model has three main applied implications for the nature of asset management companies and their role in corporate governance.

Which asset managers will monitor. Our analysis has implications for the degree to which different types of asset managers should be expected to engage in the monitoring of portfolio firms. In particular, we show that asset managers' (i.e., FMs') skin in the game, which determines their level of monitoring, is increasing in the endowment of each client (FC) in the fund. Thus, if FCs have relatively high endowments, they will invest in funds in which managers take larger personal stakes and monitor aggressively. If, on the other hand, FCs have relatively low endowments, they will invest in funds in which managers will take small personal stakes and monitor very little.

This depiction resonates with key characteristics of asset management firms observed in reality. Relatively poor real-world investors tend to invest in mutual funds. It is well documented that mutual fund managers invest very little in their funds (Khorana, Servaes, and Wedge 2007), and mutual funds are notorious for being muted in their engagement efforts (e.g., Bebchuck et al 2017). In contrast, wealthy individuals tend to invest in hedge funds. Managers of these funds are well known to self-invest significantly (Agarwal, Daniel, and Naik 2009) and often play an active role in the monitoring of their portfolio firms (Brav, Jiang, and Kim 2010).

Large blocks may monitor less than small blocks. Our results imply that block size may not be a good predictor of monitoring intensity. With proprietary blocks as in APZ, larger stakes imply more monitoring because stake size directly determines monitoring intensity. However, with delegated blocks the fund's internal incentive structure separates monitoring incentives from stake size. In particular, in our model the endogenous block size is increasing in both the number of FCs and their initial endowment, whereas monitoring intensity is determined only by their initial endowment. As a result, blocks held by funds with many investors with low initial endowment may be larger but feature significantly less monitoring than those held by funds with a smaller number of investors with higher initial endowment. In this regard, our results are consistent with Nockher (2022), who shows that smaller blockholders tend to be more intensive monitors than larger blockholders.

The need for index funds in governance. At the broadest level, our analysis indirectly highlights a role for index funds in corporate governance. Our analysis speaks to the concentrated holding choices of *active* funds, who make deliberate portfolio decisions (as our FMs do). A key implication of our model is that such active asset managers do not utilize their full risk-bearing capacity to hold concentrated positions, and expend suboptimally low levels of resources into monitoring. This finding must be viewed in the context of the evolution of the asset management industry and the emergence of passive, i.e., index, funds which—purely by virtue of their size—mechanically end up holding concentrated positions in firms. If active funds do not hold sufficiently concentrated stakes and limit their monitoring, as our results suggest, it becomes all the more important to understand the role of index funds in governance (Brav, Malenko, and Malenko 2022, Corum, Malenko, and Malenko 2023).

1.1 Related literature

Our paper relates most directly to APZ and papers that generalize and extend their findings. DeMarzo and Urosevic (2006) extend APZ to a fully dynamic setting, while Marinovic and Varas (2021) also incorporate private information. In contrast to these papers, which retain APZ's focus on proprietary blocks, we focus on the economics of delegated block ownership rather than on dynamics.

More broadly, our work relates to a number of different literatures. At the most basic level, our paper is connected to the significant theoretical literature that studies blockholder monitoring. This literature is surveyed by Edmans and Holderness (2017). Several papers within this literature, starting with Shleifer and Vishny (1986), take block size as given while others, starting with Kyle and Vila (1991), use market microstructure frictions to model the endogenous emergence of blocks as a result of the existence of short-term trading profits. Our analysis differs from all these prior papers by explicitly modeling the emergence of *delegated* equity blocks. Further, given that our interest is in long-term block formation, we assume fully transparent financial markets, so there are no trading profits. In their analysis of optimal ownership structure, Bolton and von Thadden (1998), also assume fully transparent financial markets but—unlike us—focus purely on proprietary blocks.

More recently, a growing theoretical literature takes the delegated nature of equity ownership seriously, and considers the role of the incentives of asset managers in corporate governance. While several papers within that literature (e.g., Dasgupta and Piacentino 2015, Piacentino 2019, Dasgupta and Maug 2023) have highlighted the negative implications of agency frictions arising from the delegation of portfolio management on the level of monitoring at portfolio firms, none of those papers endogenize the emergence of delegated blockholders or the size of their blocks.

Finally, our paper is related in spirit to the literature on the endogenous emergence of financial intermediaries, starting with the work of Diamond and Dybvig (1983), as well as the literature on optimal contracting in delegated portfolio management, starting with the work of Bhattacharya and Pfleiderer (1985). Relative to the former, which has focused on banking, we consider the emergence of asset managers. Relative to the latter, which considers optimal contracting with respect to trading by asset managers, we incorporate monitoring considerations as well.

2 Proprietary blocks: A benchmark model

We start with a simplified, benchmark, version of the APZ model. Consider a financial market with a single firm that has one infinitely divisible equity share outstanding, and a risk-free asset in perfectly elastic supply whose gross return is normalized to unity. There is a unit continuum of investors who have CARA utility, each with risk tolerance of ρ. To mirror the assumption of an exogenously specified large investor in APZ, we assume that a measure $\lambda < \frac{1}{2}$ of such investors are exogenously aggregated into a single entity, L, who trades strategically taking her price impact into account, and can monitor the firm to improve its cash flows. The remaining $1 - \lambda$ of infinitesimal or "small" investors act perfectly competitively. We assume that L has an endowment of $\omega \in (0, \lambda]$ shares while the remaining $1 - \lambda$ investors have an aggregate endowment of $1 - \omega$ shares, shared equally among them.

There are three dates. Potentially numerous rounds of trading opportunities are available at date 1 in a Walrasian market: in any given round of trade, investors submit demand functions and a market-clearing price is determined. At date 2, L can choose to monitor the firm as follows: at a cost of $c(m)$, where $c'(\cdot) > 0$ and $c''(\cdot) > 0$, she can exert monitoring effort $m \geq 0$ to generate a final equity payoff that is distributed according to $N\left(\mu(m), \sigma^2\right)$, where $\mu'(\cdot) > 0$ and $\mu''(\cdot) \leq 0$. At date 3, all payoffs are publicly realized. As in the bulk of the APZ analysis, L cannot commit to a final round of trade at date 1 or to a particular level of monitoring at date 2.[4]

Aggregate risk tolerance. In a Walrasian CARA-Normal market with symmetric information like ours, at any price P, each competitive investor will have a demand of $\rho \frac{\mu(m)-P}{\sigma^2}$, and thus the total demand of a measure x of small competitive investors is given by $\rho x \frac{\mu(m)-P}{\sigma^2}$, which is equivalent to the demand of a *single* competitive investor with risk tolerance of ρx. In other words, the *aggregate risk tolerance* of a given measure of small competitive investors is proportional to the measure of those investors. Accordingly, we treat the $1 - \lambda$ measure

[4] APZ also consider the case of multiple assets as well as more general monitoring technologies; we use this baseline version of their model, as it is under these specific assumptions that APZ provide the most complete characterization for us to benchmark against.

of infinitesimal investors as being represented by a single competitive investor with risk tolerance $\rho(1-\lambda)$. For benchmarking purposes, we assume that L has the same risk tolerance as the aggregate risk tolerance of the measure of competitive investors he replaces, i.e., $\rho\lambda$. This assumption will be convenient when we generalize the model to explicitly model the large investor as an endogenous delegated trading vehicle, i.e., a fund.

Competitive allocations with perfect risk sharing. Before analyzing the full equilibrium involving both strategic and competitive trading as well as monitoring, it is helpful to establish a benchmark in which all investors are competitive and monitoring cannot arise. In such a benchmark, risk sharing considerations are the sole determinants of equilibrium allocations. Denoting $L's$ equilibrium holdings by α, it is easy to see that the *competitive equilibrium allocation* is $\alpha = \lambda$. This is because the competitive equilibrium involves perfect risk sharing, under which L would hold $\frac{\rho\lambda}{\rho\lambda+\rho(1-\lambda)} = \lambda$ fraction of the risky asset while the small investors would hold $\frac{\rho(1-\lambda)}{\rho\lambda+\rho(1-\lambda)} = 1-\lambda$ of the risky asset, in accordance with their relative levels of risk tolerance.

Equilibrium trading and monitoring. Given that L is unable to commit to a particular level of monitoring, her monitoring is determined by her final holdings on date 2. If α is $L's$ total ownership of the risky asset upon entering date 2, then her equilibrium monitoring level is given by $m(\alpha) = argmax_m \Psi(\alpha)$, where

$$\Psi(\alpha) = \alpha\mu(m) - c(m) - \frac{1}{2\rho\lambda}\alpha^2\sigma^2 \tag{1}$$

is the certainty equivalent for L of holding α units of the risky asset and monitoring at intensity m. The optimal level of monitoring is given implicitly by

$$\alpha = \frac{c'(m(\alpha))}{\mu'(m(\alpha))}. \tag{2}$$

Clearly, $m(\alpha)$ is increasing in α.

If $L's$ final ownership of the risky asset is expected to be α, the $1-\lambda$ measure of small investors have an aggregate demand of $\rho\,(1-\lambda)\,\frac{\mu(m(\alpha))-P}{\sigma^2}$, giving rise to a market clearing price

$$P\,(\alpha) = \mu(m(\alpha)) - \frac{1-\alpha}{\rho(1-\lambda)}\sigma^2. \tag{3}$$

Finally, given that L is unable to commit to a final round of trade within date 1, we follow APZ in focusing on *globally stable allocations.* Such an allocation is defined as follows:

Definition 1. An allocation α_G is globally stable iff (i)

$$\alpha_G \in argmax_\alpha \Psi(\alpha) - \Psi(\alpha_G) - (\alpha - \alpha_G)P(\alpha_G),$$

and (ii) for every $\omega \in [0,1]$, such that $\omega \neq \alpha_G$,

$$\Psi(\alpha_G) - \Psi(\omega) - (\alpha_G - \omega)P(\alpha_G) > 0.$$

In words, this means that: (i) once a globally stable allocation is reached, L will not wish to trade away from it at current prices; and (ii) L is willing to trade to the globally stable allocation from any other position at prices consistent with the globally stable allocation. In their central result, APZ show that:

Proposition 1. (Admati, Pfleiderer, and Zechner 1994, Proposition 4) *As long as $\Psi(\cdot)$ is strictly concave, there exists a unique globally stable allocation, $\alpha_G = \lambda$, which coincides with the competitive equilibrium allocation.*

All proofs are in the Appendix. This key result implies that the possibility of monitoring does not affect the degree of risk sharing in equilibrium. The reason is that the lack of ability to commit to a final round of trade erodes the strategic advantage of the large investor, who subsequently trades all the way to the competitive equilibrium allocation. Put another way, an endowment effect induces L to trade all the way to the risk sharing optimum. Counterfactually, if any sequence of trades led to a proposed final holding level for L that is strictly below her risk sharing optimum, she would be tempted to buy a bit more because the current holding is now part of her endowment. Starting from this endowment, there will always be at least some incremental risk sharing gains by buying a bit more, despite having to pay the full value of future monitoring in making such purchases.

While the concept of global stability is static, its relevance has been confirmed by DeMarzo and Urosevic (2006) in a fully dynamic version of the APZ model with continuous trading and monitoring opportunities. Indeed, their main result is that the large trader will ultimately trade to the competitive price-taking allocation, which generalizes and provides a dynamic micro-foundation for APZ's concept of global stability.

It is useful to note that the APZ equilibrium is clearly inferior to a social planner's optimum. While risk sharing is optimal, the unconstrained optimum would have monitoring determined as if the large investor owns the entire firm, so monitoring is clearly suboptimally low relative to the first best. A full welfare analysis is provided in section 3.2 below.

Overall, APZ conclude that proprietary investors with high risk-bearing capacity will indeed acquire large blocks and thus monitor intensively. In reality, the concentration of ownership in the hands of a large investor is typically achieved by delegating portfolio management to professional asset managers who trade and monitor on behalf of their clients. Thus, in the remainder of the paper, we examine how incentives to monitor are determined when risk averse investors can optimally delegate to professional fund managers, who can then trade freely in financial markets but cannot make prior commitments to monitor firms at any particular level of intensity.

3 Delegated blocks

Most equity blocks today are held by delegated investment funds, in which professional fund managers manage capital for non-professional investors.[5] To model blocks held by funds, we assume that a measure $1-\lambda$ of agents are "financial" investors while the remaining λ measure of agents are "nonfinancial" investors. Financial investors are professionals who specialize in trading and monitoring. They can also provide such services to nonfinancial investors. As a group, they are endowed with $1-\omega$ shares, distributed equally among themselves. Non-

[5] Bolton and Rosenthal (2019) document a transition from proprietary blocks to delegated blocks over time for an important class of US firms.

financial investors do not participate directly in financial markets. This is because they are otherwise occupied and face a high opportunity cost for trading or monitoring. In section 4.2, we derive a lower bound on the opportunity cost that supports our results. As a group, nonfinancial investors are endowed with $\omega \in (0, \lambda)$ shares, distributed equally among themselves.[6] Nonfinancial investors thus have a lower endowment of the risky asset than their desired holdings in a competitive equilibrium, and—given the opportunity to participate in financial markets by joining funds—could enjoy risk sharing benefits by increasing their risky asset exposure.

To operationalize the idea of fund formation, we introduce an outside agent (who does not otherwise participate in the game) who proposes contracts, i.e., a "fund designer." A fund is formed when the fund designer offers a contract to a measure of nonfinancial investors—who become Fund Clients (FCs) if they agree to join—and a measure of financial investors—who become Fund Managers (FMs) if they agree to join.[7] Investors, whether financial or nonfinancial, agree to join the fund only if their incentive compatibility conditions are satisfied. The contract requires that the fund become the sole trading vehicle for both FMs and FCs, who coinvest in the fund by contributing their initial endowments as well as possible cash injections. The FMs make trading and monitoring decisions on behalf of the fund based on their contractual incentives. As in the APZ benchmark, the FMs cannot commit to a given trading strategy or monitoring level up front—they always behave opportunistically once the fund has been established.

Our model captures key aspects of real world financial markets. In such markets there are professional asset managers, indirect investors who invest only via those professional managers, and investors who trade directly on their own account. In our model, professional asset managers are represented by FMs, indirect investors are represented by FCs, and those who trade directly are represented by financial investors who do not become FMs.

[6] The assumed equality between the measure of investors represented by L in the baseline model and the measure of nonfinancial investors in the main model is purely for expositional convenience. All our qualitative results hold for any $\lambda \in (0, 1)$.

[7] We show later that no funds would form without the participation of nonfinancial investors. See section 4.3 for a full discussion.

Optimal delegation. We have a principal-agent model in which the nonfinancial investors are the principals and the financial investors are agents. For the majority of our analysis, we take the conventional approach in which we solve for the optimal fund as the contract that maximizes the payoffs of the principal, i.e., the nonfinancial investors. Accordingly, we assume that the fund designer proposes a contract that maximizes the payoffs of the nonfinacial investors as a whole, which implies that all nonfinancial investors will be FCs. In section 5.1 we solve the complementary case in which the fund designer proposes a contract that maximizes the profits of the fund.

We conduct our baseline optimal contracting analysis in two steps. First, we find the optimal allocation from the point of view of the FCs as if they could act as a single strategic financial investor (like L in section 2), and in addition have the ability to commit ex ante to both a given monitoring level and a single round of trade. This corresponds to their optimal payoff with full commitment ability vis a vis both monitoring and trading. We then show that under intuitive and plausible conditions a simple delegated fund contract exists that fully achieves this optimal outcome for the FCs, subject to the FMs' actual trading and monitoring decisions under the contract. In other words, while no investors in the model actually have the ability to commit to a monitoring level or a trading strategy, we show that optimal delegation can achieve the full-commitment optimum for the FCs.

If the FCs could act as a strategic financial investor and publicly commit to a monitoring level of m and a single round of trade, the price they would face if they traded to a final stake of α is given by $\mu(m) - \frac{1-\alpha}{\rho(1-\lambda)}\sigma^2$. Thus, they would have a joint optimization problem given by

$$\max_{m,\alpha} \alpha\mu(m) - c(m) - \frac{1}{2\rho\lambda}\alpha^2\sigma^2 - (\alpha - \omega)\left(\mu(m) - \frac{1-\alpha}{\rho(1-\lambda)}\sigma^2\right).$$

Solving this problem yields the following result.

Proposition 2. *The FCs' full-commitment optimum has an optimal monitoring level, m^C, implicitly defined by $\omega = \frac{c'(m^C)}{\mu'(m^C)}$, and an optimal final stake of $\alpha^C \equiv \frac{\lambda(1+\omega)}{(1+\lambda)}$.*

Since $\omega < \lambda$, it is easy to see that the optimal final stake of the FCs lies between their initial endowment (ω) and their competitive allocation with perfect risk sharing (λ):

$$\omega < \alpha^C < \lambda.$$

The FCs want to increase their stake in the risky asset above their endowment to enhance their risk sharing (so $\alpha^C > \omega$ is optimal). However, they avoid trading all the way to their competitive allocation so that they can fully exploit their strategic trading advantage, i.e., accounting for the fact that they move prices (so $\alpha^C < \lambda$). Further, the FCs' full-commitment optimal monitoring level does not depend on their final stake, α. Instead, by analogy to equation (2), it is clear that the FCs desire monitoring to occur "as if" their ownership was equal to their initial endowment ω. Thus, FCs want to hold more than their initial endowment for risk sharing purposes but wish to monitor only at their original endowment level. Intuitively, this is because FCs are aware that any increase in monitoring over and above the level implied by their endowment induces a higher price that offsets the benefits of the additional monitoring from their perspective. Let

$$\Pi_{FC}^C \equiv \alpha^C \mu(m^C) - c(m^C) - \frac{1}{2\rho\lambda}(\alpha^C)^2\sigma^2 - (\alpha^C - \omega)\left(\mu(m^C) - \frac{1-\alpha^C}{\rho(1-\lambda)}\sigma^2\right)$$

denote the FCs' aggregate equilibrium payoff at their full-commitment optimum.

Fund formation. We now identify conditions under which it is possible to replicate the payoff Π_{FC}^C for the FCs by forming a single fund that all nonfinancial investors are willing to join. At the outset, it is important to note that even if a fund contract can be designed that delivers exactly this optimal payoff to the FCs, it may not be feasible to form such a fund because individual nonfinancial investors will not be willing to join it.

Lemma 1. *There exists $\hat{\omega} \in (0, \lambda)$ such that for $\omega \leq \hat{\omega}$, nonfinancial investors will join a fund that delivers an aggregate FC payoff of Π_{FC}^C, while for $\omega > \hat{\omega}$ they will not.*

In other words, delegated blockholding can arise only when nonfinancial investors have relatively low endowments of the risky asset. Intuitively, if an individual FC defects from the fund, they enjoy the full benefits of the fund's monitoring for free and lose only the risk sharing benefits that arise from accessing financial markets through the fund. Thus, if the endowment ω was sufficiently close to the competitive risk sharing level, and risk sharing benefits were therefore small, individual FCs would prefer to defect from the fund.

Now consider a proposed fund contract specified as follows: a chosen mass of FMs, $\tau \in (0, 1-\lambda)$, invited into the fund, a skin in the game parameter, $\phi \in [0,1]$, specifying the FMs' share of the fund's assets, and an up-front fee, f, which each FC must pay to join the fund. Overall, a fund formed under this contract can be represented as a contracting triple, (τ, ϕ, f), representing the measure of FMs, their skin in the game, and the per-FC fee, respectively.

Since the FMs undertake trading and monitoring on behalf of the fund once it is formed, decisions in the fund are taken by a measure τ of investors who have a final ownership stake of $\phi\alpha^D$, where α^D is the final overall holding of the fund. As discussed in section 2, with respect to trading, if each FM trades according to their individual risk tolerance, their aggregate trade can be represented as that of a single competitive investor with risk tolerance $\rho\tau$. In other words, with respect to trading, it does not matter whether we assume that all FMs act individually or that they act jointly as a single competitive investor with the appropriate risk tolerance. The situation with respect to monitoring is more involved. As members of a continuum, if FMs behaved individually with respect to monitoring, complete free-riding would arise and no monitoring would occur. To give monitoring a chance, we thus assume that FM's make monitoring decisions as a group based on their aggregate contractual incentives once they join the fund, i.e., FMs choose the level of monitoring as would a single investor with a final block of size $\phi\alpha^D$.[8] That said, there are three important points to note. First, the decision to join the fund must still be incentive compatible for each financial investor *individually*. Second, assuming the ability of FMs to behave as a group with respect to monitoring once inside the fund does *not* imply that monitoring must arise. We show in section 5.1 that despite such group behavior there can be funds with no monitoring. Finally,

[8] There are multiple ways to microfound this. For example, it could be that the measure τ of FMs, after joining the fund, forms a group that assigns monitoring duties to its members such that total monitoring is optimal for a single investor who owns a block of size $\phi\alpha^D$, and is able to perfectly enforce the compliance of all members of the group. Alternatively, a single FM could be selected and asked to monitor at a level corresponding to a holding of $\phi\alpha^D$, while being compensated by other FMs for her cost in doing so.

group behavior cannot generate monitoring by funds involving *only* financial investors, as we show in section 4.3.

We solve for the optimal fund by backward induction. We first assume that a fund with λ FCs and (an arbitrary positive measure) τ FMs is formed, and proceed to compute the monitoring and trading decisions of the FMs for a given (τ, ϕ, f). We then solve for the optimal contracting terms that achieve a payoff of Π_{FC}^{C} for the FCs. While doing so, we ensure that all τ FMs are willing to join the fund. We denote the optimal set of contracting terms by the triple (τ^*, ϕ^*, f^*).

If α^D is the fund's total ownership of the risky asset upon entering date 2, then FMs hold an *effective stake* of $\phi\alpha^D$. FM's will optimally choose their monitoring intensity to solve: $m^D(\alpha^D) = argmax_m \Psi^D(\alpha^D)$ where

$$\Psi^D(\alpha^D) = \phi\alpha^D\mu(m^D) - c(m^D) - \frac{1}{2\rho\tau}\phi^2\left(\alpha^D\right)^2\sigma^2 \tag{4}$$

is the certainty equivalent for the FMs if they hold an effective stake of $\phi\alpha^D$ units of the risky asset and monitor at intensity m^D. The solution to the above optimization problem is given implicitly by $\phi\alpha^D = \frac{c'(m^D(\alpha^D))}{\mu'(m^D(\alpha^D))}$.

Since the FMs cannot commit to a given trading strategy, we again focus on globally stable trading allocations. We note first that the pricing function must be adjusted for the fact that the mass of competitive price-taking investors has been reduced from $1-\lambda$ to $1-\lambda-\tau$ given the formation of the fund. If the competitive investors expect the fund to end up with a stake of α^D, their aggregate demand will be $\rho\left(1-\lambda-\tau\right)\frac{\mu(m^D(\alpha^D))-P}{\sigma^2}$, giving rise to a market clearing price of

$$P^D\left(\alpha^D\right) = \mu(m^D(\alpha^D)) - \frac{1-\alpha^D}{\rho(1-\lambda-\tau)}\sigma^2. \tag{5}$$

The definition of a globally stable allocation must also be adjusted for our delegated fund model since FMs make decisions on behalf of the entire fund but enjoy only a ϕ proportion of its payoff.

Definition 2. An allocation α_G^D is globally stable iff (i)

$$\alpha_G^D \in argmax_\alpha \Psi^D(\alpha) - \Psi^D(\alpha_G^D) - \phi(\alpha - \alpha_G^D)P^D(\alpha_G^D),$$

and (ii) for every $\omega \in [0, 1]$, such that $\omega \neq \alpha_G^D$,

$$\Psi^D(\alpha_G^D) - \Psi^D(\omega) - \phi(\alpha_G^D - \omega)P^D(\alpha_G^D) > 0,$$

We have the following result.

Lemma 2. *As long as $\Psi^D(\cdot)$ is strictly concave, there exists a unique globally stable allocation*

$$\alpha_G^D = \frac{\tau/\phi}{\tau/\phi + 1 - \lambda - \tau}. \tag{6}$$

The equilibrium stake of the fund, α_G^D, depends only on the sizes of the FC and FM populations, λ and τ, and the skin in the game parameter, ϕ. The expression is analogous to the globally stable allocation in the APZ benchmark derived in Proposition 1: it is part of the competitive equilibrium allocation in a market with two competitive traders, one of whom has risk tolerance of $\rho(1 - \lambda - \tau)$ while the other has risk tolerance of $\rho\tau/\phi$. The former is simply the aggregation of the financial investors who did not join the fund (i.e., did not become FMs). The latter is the fund itself, which can be represented by an investor with risk tolerance of $\rho\tau/\phi$. This is because, as discussed above, the measure τ of FMs (with aggregate risk tolerance of $\rho\tau$) who make trading decisions are only exposed to a fraction ϕ of the holdings of the fund.

Notably, however, this allocation does not necessarily correspond to perfect risk sharing among all investors—which arose in the globally stable allocation of APZ (see Proposition 1)—as this would require an allocation of $\alpha_G^D = \tau + \lambda$ (since the measure of investors in the fund is the sum of λ FCs and τ FMs). The deviation from perfect risk sharing arises due a combination of two factors: First, only a measure of $\tau < \tau + \lambda$ investors make decisions on behalf of the whole fund; and second, those investors are exposed to only a fraction ϕ of the fund's holdings. Indeed, it is apparent that if $\tau/\phi = \tau + \lambda$ in the expression for α_G^D above, we obtain $\alpha_G^D = \tau + \lambda$.

We assume for the remainder of the analysis that $\Psi^D(\cdot)$ is strictly concave, so that a globally stable allocation exists. Comparing the effective stakes of financial investors inside and outside the fund yields the following result.

Lemma 3. *The FMs in the fund and the outside financial investors end up with identical effective per-investor holdings of the risky asset.*

This result implies that there is perfect risk sharing among the total $1 - \lambda$ measure of financial investors, whether inside or outside the fund, over the part of the risky asset that is not (effectively) held by the FCs. This is because the existence of multiple trading opportunities combined with the inability to commit to a particular trading strategy erodes the strategic advantage of the FMs, who subsequently trade to arrive at the point of perfect risk sharing between themselves and financial investors outside the fund.

Given the trading and implied monitoring choices of the FMs for a given (τ, ϕ, f), we now proceed to solve for the optimal contracting terms to determine (τ^*, ϕ^*, f^*).

Proposition 3. *For $\omega \leq \hat{\omega}$, it is feasible to form an optimal fund that achieves an aggregate payoff of Π_{FC}^C for the FCs. It is characterized by:*

1. *a mass of FMs $\tau^* = \frac{(1-\lambda^2)\omega}{1-\lambda\omega}$,*
2. *a skin in the game parameter $\phi^* = \frac{(1+\lambda)\omega}{2\lambda\omega+\lambda+\omega}$, and*
3. *a fee*

$$f^* = \frac{1}{\lambda}\left[c(m^C) + P^{D*}(\alpha_G^{D*})\left((1-\phi^*)(\omega + \tau^*\frac{(1-\omega)}{1-\lambda}) - \omega\right)\right] \tag{7}$$

where the superscript $D*$ indicates that the associated function or variable is evaluated at ϕ^* and τ^*.

The proof proceeds in several steps. First, we solve for τ^* and ϕ^* such that (1) the FMs' equilibrium monitoring effort equals the FCs' optimal level, m^C—which requires that the FMs' effective stake in the risky asset, $\phi^*\alpha_G^{D*}$, equals ω—and (2) that the FCs' effective stake, $(1-\phi^*)\alpha_G^{D*}$, equals their optimal stake, α^C. Next, we set the fee f^* to just satisfy the participation constraint of individual FMs conditional on the existence of a fund involving τ^* FMs with a skin in the game parameter ϕ^*. Finally, we show that this fee level makes the FCs' aggregate payoff coincide exactly with Π_{FC}^C, which also ensures that each FC will participate given that $\omega \leq \hat{\omega}$.

We examine the implications of this result and the intuition behind it through a series of remarks. First, we discuss the levels of risk sharing and monitoring implied by the equilibrium. The optimal delegation contract selects the skin in the game parameter (ϕ^*) and measure of FMs (τ^*) so as to render the effective holdings of the FMs equal to the endowment of the FCs (ω). FCs, however, do *not* hold their original endowment, but rather end up with a larger holding of the risky asset: $\lambda\frac{1+\omega}{1+\lambda} > \omega$. As explained previously, these outcomes correspond to what the FCs would like to do if they could (counterfactually) exert full commitment power over their monitoring and trading strategies. With respect to monitoring, FCs prefer that monitoring occur at a level commensurate with their initial endowment due to the well-known free-riding phenomenon (first formalized by Grossman and Hart 1980). Effectively, the FCs cannot "enjoy" monitoring benefits on anything more than their initial endowment, and thus would like to limit it. However, to obtain risk exposure they would like to hold a larger stake than their initial endowment. In the absence of commitment power, doing these things simultaneously is not possible in a proprietary blockholder model like APZ, where the final stake of the blockholder drives both monitoring incentives and risk sharing benefits. Thus, the key question is: how is this accomplished in our model, which still does not give investors any commitment power?

The answer is that, in our model with delegation, FCs do not directly monitor—FMs monitor on their behalf. That is, delegation by definition splits roles: FCs own a fraction of the fund, $(1-\phi^*)\,\alpha_G^{D*}$, but monitoring is undertaken by the FMs who own the rest: $\phi^*\alpha_G^{D*}$. Thus, delegation enables the separation of (FC-) ownership and (FM-) monitoring. The contract that maximizes FC welfare ensures that monitoring occurs only at the level that is privately optimal for FCs absent risk sharing considerations, but gives them an effective stake in the fund that also optimizes their collective risk sharing and trading incentives without regard to any effect on monitoring. To summarize:

Remark 1. The delegation of blockholding breaks the link between risk sharing by FCs and the monitoring that occurs on their behalf. The optimal contract enables monitoring at a level privately optimal for FCs absent risk sharing motives, while also enabling FCs to trade to their privately optimal level of risk sharing absent monitoring motives, taking into account their collective market power.

Put differently, outcomes in our model are affected by the ability to write a contract, which therefore has the flavor of a commitment device. However, it is worth noting that we do *not* require long-term commitment to fund contracts. A full discussion of the robustness of our results to the possibility of recontracting is provided in section 4.1.

Second, we consider the role of the fee, f^*. The fee is relevant to both FMs who receive it and FCs who pay it. We consider FMs first. Any individual financial investor can choose between joining the measure τ^* of FMs inside the fund or remaining part of the measure $1-\lambda-\tau^*$ of financial investors who do not join the fund. As noted above, the fee f^* is set to make the payoffs of these two choices equal.

Since, as shown in Lemma 3, FMs in the fund and financial investors outside the fund end up holding the same effective stake per investor, their degree of risk sharing is not affected by joining the fund. Thus, the FCs do not need to compensate the FMs for taking more or less risk. However, there are two ways in which FMs' payoffs differ from those of financial investors who choose not to become FMs. First, financial investors choosing to become FMs share the cost of monitoring, while those remaining outside the fund enjoy the benefits of such monitoring for free. Thus, the fee f^* must compensate FMs for their monitoring costs at the fund's ultimate stake. Second, the fee must compensate FMs for giving up some of their endowment. When they join the fund, each FM is allocated an initial pre-trade endowment of $\frac{\phi^*}{\tau^*}\left(\omega+\tau^*\frac{1-\omega}{1-\lambda}\right)$, because each FM gets a $\frac{\phi^*}{\tau^*}$ share of the fund, to which the FCs jointly contribute ω and FMs jointly contribute $\tau^*\frac{1-\omega}{1-\lambda}$. This pre-trade endowment $\frac{\phi^*}{\tau^*}\left(\omega+\tau^*\frac{1-\omega}{1-\lambda}\right)$ is smaller than $\frac{1-\omega}{1-\lambda}$, the initial endowment of each financial investor.

Remark 2. The fee f^* compensates FMs for their expected equilibrium monitoring costs as well as for the value of any endowment that they sacrifice when they join the fund.

Next, we turn to the role of the fee in the FCs' payoff. Clearly, if FMs require compensation for monitoring costs incurred inside the fund then FCs must pay for such costs. This is consistent with the FCs' full-commitment optimum, in which they also (counterfactually) pay the full monitoring cost. Further, in contrast to FMs, FCs start with an endowment in the fund that is *higher* than their initial endowment of $\frac{\omega}{\lambda}$. In line with the discussion above, this is a result of the reallocation of some of the FMs' initial endowment to the FCs. FCs must pay for this added initial endowment. Put another way, in order to achieve their full-commitment optimum level of ownership, α^C, FCs effectively have to buy less as a result of joining the optimal fund than they would if, counterfactually, they traded independently to this allocation. The fee f^* charges them for this benefit, thus bringing their total payoff to Π_{FC}^C.

Remark 3. The fee f^* charges the FCs for the full anticipated monitoring costs expended by FMs as well as as for the value of any additional endowment that they are allocated when joining the fund.

3.1 Risk sharing and monitoring: Delegated vs proprietary ownership

We are now in a position to compare our results on delegated blockholding to those of APZ's benchmark (presented in Section 2). Taking as given the existence of a proprietary trader with large risk-bearing capacity, APZ ask whether the anticipation of monitoring costs affects the degree of risk sharing in the economy. Under broad and plausible conditions, they find the answer is "no"—the proprietary trader will still trade to the competitive risk sharing allocation and monitor at that allocation. However, we show that when blockholding is achieved by optimal delegation, the picture changes dramatically, in at least two ways.

First, the delegated vehicle that is formed holds *less* of the risky asset, i.e., a smaller block, than what is implied by perfect risk sharing, whereas in the proprietary APZ case perfect risk sharing is achieved. To see this formally, first note that the optimal fund holds a stake in the risky asset equal to $\omega + \lambda\frac{1+\omega}{1+\lambda}$, i.e., the effective stake of the FMs plus the effective stake of the FCs. It is straightforward to show that this is less than the fund's competitive risk sharing optimal allocation, $\lambda + \tau^*$, which is also what a proprietary blockholder representing the same measure of traders would hold in a globally stable allocation.

Corollary 1. *The optimal fund holds strictly less of the risky asset than the corresponding competitive equilibrium allocation for a trader with the same overall risk tolerance.*

Intuitively, when delegating to form a fund, the optimal contractual terms account for the fact that the fund will affect prices when trading and thus ensures that the FCs ultimately hold an amount of the risky asset that reflects their market power (and thus the fund shades its trades downwards). Thus, in comparison to the APZ proprietary benchmark, delegated blockholding results in suboptimal risk sharing.

Second, delegation separates ownership and monitoring by allocating monitoring tasks only to a subset of participants in the fund, i.e., the FMs. The optimal delegation contract allocates an effective stake for FMs of ω, which results in a level of monitoring that would be privately optimal for FCs absent risk sharing considerations, i.e., corresponding purely to their initial endowments (see Remark 1). As a result:

Corollary 2. *The optimal fund undertakes strictly less monitoring than a proprietary blockholder would if they held a block of identical size.*

3.2 Welfare comparisons

We now conduct a full welfare analysis. We compare our delegated blocks equilibrium to APZ's benchmark, the first best social planner's optimum, and an "autarky" setting in which nonfinancial investors are unable to join funds and must hold their endowment. To do so, it is helpful to define for each setting two variables that index the efficiency of monitoring and the efficiency of risk sharing. The monitoring level depends directly on the effective stake of the monitoring entity—the large shareholder in the APZ benchmark, the FMs in our delegated blocks model, and any subset of investors chosen by the planner in the first best. Thus, we let α_M^S index the monitor's effective stake for setting $S \in \{FB, APZ, DB, A\}$, where FB represents the first best planner's solution in which the planner can fully dictate investors' final holdings and their monitoring choices, APZ is the APZ benchmark, DB is our delegated blocks equilibrium, and A is autarky (in which there is no monitoring entity). Similarly, with respect to risk sharing, since in all of these settings the financial investors end up with identical holdings of the risky asset, we can index the degree of risk sharing with the effective stake of the λ-sized group of investors represented either by L (in the APZ model) or the FCs (in our model), which we denote by α_R^S for each setting.

Given these definitions, it is easy to see that the first best has $\alpha_M^{FB} = 1$ and $\alpha_R^{FB} = \lambda$ so that monitoring and risk sharing are both optimal.[9] In the APZ benchmark the effective stake of L determines both the level of monitoring and the efficiency of risk sharing, so $\alpha_M^{APZ} = \alpha_R^{APZ} = \lambda$. In our delegated blocks model, monitoring is determined by the effective stake of the FMs, while the efficiency of risk sharing depends on the effective stake of the FCs. We thus have $\alpha_M^{DB} = \omega < \lambda$ and $\alpha_R^{DB} = \alpha^C = \lambda \frac{1+\omega}{1+\lambda} < \lambda$. This gives us a clear welfare ranking: the APZ benchmark is inferior to the first best because of suboptimal monitoring, and our delegated blocks equilibrium is inferior to the APZ benchmark because it has even lower levels of monitoring as well as inferior risk sharing.

Finally, consider the autarky setting in our model. If nonfinancial investors are forced to hold their endowment, and no fund is formed so there is no monitoring, we have $\alpha_M^A = 0$ and $\alpha_R^A = \omega < \alpha^C = \lambda \frac{1+\omega}{1+\lambda}$. Thus, this setting has the lowest overall welfare. To conclude, our delegated blocks equilibrium is clearly inferior to the proprietary blocks benchmark of APZ, but it does provide improvements in both monitoring and risk sharing relative to autarky.

4 Discussion and robustness

4.1 Recontracting

In APZ, there is no trade-off between risk sharing and monitoring, because an endowment effect induces the large trader, L, to trade all the way to the risk sharing optimum. Start-

[9] Note that the planner can assign these monitoring duties to any subset of investors with monitoring skill, and can enforce transfers of asset holdings or money among investors to achieve these objectives.

ing from any stake less than the risk-sharing optimum, there will always be at least some incremental risk sharing gains by buying a bit more, despite having to pay the full value of future monitoring in making such purchases. In the optimal contract solved above, FCs end up with a stake of $\frac{\lambda}{1+\lambda}(1+\omega)$, which is greater than their initial endowment of ω, but monitoring occurs at a level corresponding to an ownership of ω. Hence, one may wonder whether a variant of the APZ endowment effect could come into play wherein the FCs now wish to recontract with FMs to reflect their new endowment. Could the possibility of such recontracting revive the APZ result, such that the FCs achieve the risk sharing optimum holding of λ, and monitoring occurs at a commensurate level?

In principle, the FCs as a group would indeed like to recontract with FMs to form a fund that monitors more intensely. To see this, consider a situation where the optimal contract from above is signed and the fund trades to the stable allocation, but then an unexpected opportunity arises *ex post* to dissolve the existing fund and start a new one prior to any monitoring taking place.[10] We then have a repeat of the model above starting from an aggregate FC endowment of $\alpha^C(\omega) = \frac{\lambda}{1+\lambda}(1+\omega)$ instead of ω, which may lead to a new fund that will monitor at a level corresponding to an ownership level of $\frac{\lambda}{1+\lambda}(1+\omega)$.[11] We refer to such recontracting opportunities as *ex post recontracting* opportunities. Could a sequence of ex post recontracting opportunities revive the APZ result? We show that this is not possible:

Proposition 4. *Ex post recontracting cannot deliver the APZ outcome.*

Intuitively, unlike repeated trading, which is always feasible, repeated contracting is constrained by free riding. As shown above in Lemma 1, as soon as FCs reach a new endowment level of $\omega > \hat{\omega}$ via some sequence of ex post recontracting opportunities, it is no longer possible to form a new optimal fund. For such endowments, the risk sharing benefits to individual FCs is too small, and thus each individual FC would benefit by deviating and staying out of the fund (if it is formed), thus saving themselves the fees that must be paid to

[10] Note that such a dissolution would necessarily involve a clawback of the part of the fee f^* that compensated the FMs for their anticipated monitoring costs.

[11] As noted in Lemma 3, FMs in the fund and financial investors outside the fund hold the same effective stakes after trading, so the new fund formation problem is isomorphic to the original problem with different endowments.

the FMs. Since $\hat{\omega} < \lambda$, and thus the FCs' ownership stake in the largest optimal fund that can be formed, $\frac{\lambda}{1+\lambda}(1+\hat{\omega})$, is smaller than λ, it is not possible to revive the APZ result by repeatedly recontracting to the optimal fund.

For $\omega > \hat{\omega}$, is it feasible to recontract to some other, suboptimal, fund with higher monitoring, and thus sequentially reach the APZ solution? This too is not feasible, because any such alternative fund offers no greater risk sharing benefit to any individual FC (because the optimal fund maximizes their individual risk sharing payoff net of price impact), but offers a higher intensity of monitoring, and thus enhances the desire to stay out of the fund. As a result, for $\omega > \hat{\omega}$, no FC would ever agree to join any fund that increases monitoring relative to the optimal fund at ω. Thus, the possibility of recontracting does not revive the APZ result.

It is also worth noting that any endowment level $\omega \leq \hat{\omega}$ for which the optimal fund can be formed but for which $\frac{\lambda}{1+\lambda}(1+\omega) > \hat{\omega}$ holds would not be subject to *any* ex post recontracting. This clearly holds for some positive measure set of endowments:

$$\Omega_S = \{\omega' > 0 : \omega' \leq \hat{\omega} < \frac{\lambda}{1+\lambda}(1+\omega') < \lambda\}.$$

4.2 Opportunity costs of market participation

We have maintained the assumption throughout that nonfinancial investors are otherwise occupied and face a high opportunity cost of participating directly in financial markets. Here we compute the opportunity cost that makes each nonfinancial investor prefer to join the optimal fund rather than pay the opportunity cost and trade on their own.

Suppose a nonfinancial investor believes that all other nonfinancial investors will join the optimal fund. Then what is the minimum cost to trade that would keep them in the fund, given that they could enjoy the fund's monitoring for free while also benefiting from risk sharing? If they pay the opportunity cost, they effectively become a financial investor who, as derived above, will trade so as to perfectly share risk with the other financial investors outside the fund (and will not monitor as they will free ride on the fund's monitoring). The minimum opportunity cost then corresponds to the difference between the payoff of a non-FM financial investor *who starts at a lower endowment* and a nonfinancial investor inside the fund. We have the following result.

Lemma 4. *The minimum opportunity cost that supports the formation of the optimal fund from Proposition 3 is given by* $\left(\frac{\lambda-\omega}{1-\lambda^2}\right)^2 \frac{\sigma^2}{2\rho} + \frac{c(m^C)}{\lambda}$.

A nonfinancial investor who pays the opportunity cost is able to enjoy the fund's monitoring for free, and thus saves the part of the per-FC fee that goes toward offsetting the FMs' monitoring cost, i.e., $\frac{c(m^C)}{\lambda}$. In addition, they will trade further than the fund would trade on their behalf, to the point where they share risk equally with all of the non-FM financial investors. This adds to the minimum opportunity cost because, as an infinitesimal trader, they can trade all the way to their risk sharing optimum (conditional on the existence of the fund) without any price impact.

4.3 Nonfinancial investors are essential for monitoring

Since we assume the ability for FMs to behave as a group inside the fund with respect to monitoring (once they have individually decided to join the fund), it is natural to wonder whether financial investors would be able to form a fund that monitors the firm without involving FCs, i.e., an FM-only fund. In other words, could group behavior by itself achieve desirable corporate governance outcomes without involving nonfinancial investors? The answer is no.

Proposition 5. *FM-only funds cannot exist.*

In any fund with only financial investors, some subset of those investors would have to monitor and thus pay costs. However, any investor that is supposed to be in the subset that monitors can choose not to join the fund, trade on their own, and enjoy exactly the same cash flow payoffs without paying the monitoring costs. So, to persuade them to join the fund, the subset that are in the fund but do not monitor must pay those that are expected to monitor. But this effectively means that monitoring costs are shared among all investors in the fund, and so the previous argument applies and individual investors who are not expected to monitor would prefer not to join the fund. Thus the presence of *other* investors in the fund willing to compensate FMs for their monitoring efforts is required, and those investors, in turn, must be willing to bear such costs even though they can unilaterally defect from the fund. Whenever FCs enjoy a sufficient risk sharing benefit from joining the fund (see Lemma 1), they fulfill this role and thus make delegated monitoring possible.

Our analysis underscores that both financial and nonfinancial investors are essential to the endogenous formation of large entities in financial markets that can undertake monitoring. Nonfinancial investors are willing to pay for asset management services that include monitoring only because of the risk sharing benefits they gain from joining funds. Financial investors, in turn, are only willing to join funds as managers and monitor because they anticipate that they will receive such fees. More generally, the formation of a vehicle that provides public goods in financial markets populated by small investors requires the existence of two distinct groups of investors: one group must be able to provide a service (here, risk sharing) that makes it worthwhile for the other group to pay for the public good (here, monitoring).

5 Extensions

5.1 Profit maximizing funds

In our main analysis we consider the optimal fund in a classic principal-agent sense, i.e., the fund designer maximizes the payoff of the principal, in our case the fund clients. In this extension we consider, instead, the economics of delegated blockholding if the fund designer maximizes the profit of the fund by making a take-it-or-leave-it offer to the nonfinancial investors. We implement this as follows. We assume that the fund designer offers a fund that maximizes the rents that can be extracted from nonfinancial investors, subject to their willingness to join the fund. Such a fund must also ensure the participation of an optimally chosen subset of financial investors who will act as fund managers.

Proposition 6. *A profit-maximizing fund designer offers a fund that undertakes no monitoring and gives fund clients the same degree of risk sharing as does the optimal fund of*

Proposition 3.

The intuition for this result is as follows. The fund designer will make FCs indifferent between joining the fund or not. The FCs cannot get risk sharing without joining the fund, so will have to pay for any incremental risk-sharing benefit from joining the fund. FCs can, however, benefit from the fund's monitoring if they stay outside the fund. As a result, they will not pay fees for monitoring per se. Yet, if the fund offers any monitoring, it must pay FMs to monitor, reducing profits. So, the profit maximizing fund designer offers a fund that only provides risk sharing and no monitoring. Our proof shows that such a fund can be implemented as the limit of a sequence of funds that offer optimal risk sharing to FCs while the measure of fund managers becomes vanishingly small.

From a social welfare perspective, a profit-maximizing fund is clearly inferior to both the optimal fund from Proposition 3 and the APZ benchmark, as it offers strictly less monitoring than either, with the same aggregate level of risk sharing as the optimal fund. However, since it offers the FCs some risk sharing benefits, it provides a welfare boost relative to autarky. Thus, our results imply that profit maximizing funds can provide a valuable risk sharing opportunity to investors, but have no value from a corporate governance perspective.

5.2 Competition among funds

To this point we have maintained the assumption that only one fund is formed, and shown that this results in markedly worse outcomes than the proprietary blocks case analyzed by APZ. It is worth exploring whether competition among funds is likely to exacerbate or ameliorate this effect. In this section we consider the possibility of perfect competition among groups of FMs, and show that in this case delegation can—in principle—lead to even worse outcomes from a corporate governance perspective. However, we also show that our benchmark results still hold in the presence of some realistic additional frictions.

Suppose that the optimal fund described in Proposition 3 is proposed, but rival FMs can propose an alternative "purely passive" fund structure to try and lure the FCs away, by offering them the same degree of risk-sharing but no in-fund monitoring, i.e., by free-riding entirely on the monitoring efforts of the optimal fund. To allow such a "purely passive fund" (PPF) we extend our definition of funds by allowing for funds that are run by a countable number of FMs, i.e., without loss of generality, by single FMs.[12] In other words, a PPF would simply be a trading vehicle that helps a defecting FC reach the same level of risk sharing as she would in the optimal fund, but charges her a strictly lower fee because there are no monitoring costs. We now characterize such a fund and check that it would, indeed, be more attractive for a defecting FC to join this fund instead of the optimal fund.

Consider, for example, a PPF designed to attract a single FC. This fund would have an initial endowment of $\frac{\omega}{\lambda} + \frac{1-\omega}{1-\lambda}$, summing the endowment of the defecting FC and the single FM, respectively. In order to offer the FC the same risk sharing as the optimal fund, the PPF would then need to offer the FC a final position of $\frac{\alpha^C}{\lambda} = \frac{1+\omega}{1+\lambda}$, where α^C is defined in Proposition 2. Thus, the FC's total cost for achieving the full-commitment level of risk-sharing via the PPF is then $\left(\frac{1+\omega}{1+\lambda} - \frac{\omega}{\lambda}\right) P^{D*}(\alpha_G^{D*})$ where $\left(\frac{1+\omega}{1+\lambda} - \frac{\omega}{\lambda}\right)$ is the total change in the FC's position and $P^{D*}(\alpha_G^{D*})$ is the equilibrium price as defined in Proposition 3 (which is the appropriate price to test this deviation since the value of the firm is determined in equilibrium by the monitoring activities of the optimal fund). In contrast, achieving the full-commitment level of risk-sharing via the optimal fund costs each FC $\left(\frac{1+\omega}{1+\lambda} - \frac{\omega}{\lambda}\right) P^{D*}(\alpha_G^{D*}) + \frac{c(m^C)}{\lambda}$, since in the optimal fund the FC must pay the FMs' monitoring cost as well. Thus by defecting to the PPF, the FC enjoys all the benefits of the optimal fund, without paying any of the monitoring cost. A defection is therefore optimal, and the equilibrium in Proposition 3 will not exist. Furthermore, an analogous deviation argument will apply to any fund with a positive amount of monitoring in equilibrium.

While such extreme free riding has the potential to severely limit the possibility of monitoring in an economy with delegated blockholding, there are several realistic frictions that might preserve fund monitoring as in Proposition 3. For example, fund formation could involve some fixed costs such as registration fees charged by a securities regulator. In that case, an equilibrium can exist featuring a fund similar to that described in Proposition 3. Indeed,

[12] Any fund with a positive measure of FMs who have a nonzero skin in the game (which would be necessary to induce them to trade) would—retaining the no commitments assumptions of the model so far—end up monitoring after the fund is formed, thus precluding a "purely passive" fund.

imagine that the formation of any fund requires a small fixed cost of $\epsilon > 0$, for which FMs would need to be compensated. Then, replacing f^* with $f^* + \frac{\epsilon}{\lambda}$, the optimal fund would survive in equilibrium as long as each FC believes that no other FC would deviate from the optimal fund, because a single infinitesimal FC would never be willing to pay the ϵ cost.

Other real-world frictions, such as non-transparent trading markets and the potential for trading profits based on asymmetric information, could also allow for the formation of funds, like our optimal fund, that monitor in equilibrium. If such trading profits are passed along to investors in the fund, these additional benefits of joining the fund could outweigh monitoring costs and overcome free riding incentives.

6 Our optimal fund and real-world asset managers

While the optimal fund described in Proposition 3 is stylized, it bears some key similarities to asset managers observed in the real world.

Fund fees are increasing in assets under management. From above, the fee is:

$$f^* = \frac{1}{\lambda}\left[c(m^C) + P^{D*}(\alpha_G^{D*})\left((1-\phi^*)(\omega + \tau^*\frac{(1-\omega)}{1-\lambda}) - \omega\right)\right],$$

where the second term represents the value of assets allocated to FCs (over and above the endowments that they contribute) when they invest in the fund, and the first represents additional payments for monitoring services provided by the FMs. In reality, fund clients use cash to invest in funds and then pay fees over and above this for asset management services, where the latter is typically an increasing function of assets under management, i.e, so called AUM fees. It is easy to see that $m^C$and the optimal fund's assets under management, $\omega + \lambda\frac{1+\omega}{1+\lambda}$, are both increasing in ω. Thus, the cash fee, $\frac{1}{\lambda}c(m^C)$, comoves with assets under management, as in reality.

Which asset managers will monitor. Our analysis of optimal delegation arrangements also has implications for the degree to which different types of asset managers engage in the monitoring of portfolio firms. Proposition 3 implies that the skin in the game of the FMs, $\phi^* = \frac{(1+\lambda)\omega}{2\lambda\omega+\lambda+\omega}$, and their total effective investment in the risky asset, ω, are both increasing in the endowment of each FC. In turn, the level of equilibrium monitoring undertaken by the fund increases in asset managers' effective stake. Thus, within the constraint under which delegated blockholding arises in equilibrium ($\omega < \hat{\omega} \in (0, \lambda)$), if fund clients have a relatively high endowment of the risky asset, they will invest in funds where managers take a larger personal stake in the fund; these funds monitor more aggressively. In contrast, if fund clients have a relatively small endowment of the risky asset, they will invest in funds where managers take a small personal stake in the fund; these funds monitor their portfolio firms very little.

This depiction of asset management resonates with key characteristics of different types of asset management firms observed in reality. Relatively poor real world investors tend to invest in mutual funds. It is well documented that mutual fund managers invest very little in their funds: according to Khorana, Servaes, and Wedge (2007) some 57% of mutual fund managers do not invest at all in their constituent funds, and the average self-investment among the rest is 0.04%. Finally, mutual funds are notorious for being relatively muted in their engagement of portfolio firms (e.g., Bebchuck, Cohen, and Hirst 2017). In contrast, relatively wealthy individuals tend to invest in hedge funds, which typically have minimal net worth requirements. Hedge funds managers are well known to self-invest significantly in the fund (estimates used in the literature range from 7% of fund assets under management in Agarwal, Daniel, and Naik 2009 to 20% in He and Krishnamurthy 2013). It is also well documented that hedge funds play a far more active role in the monitoring of their portfolio firms (Brav, Jiang, and Kim 2010).

Larger blocks may monitor less than small blocks. Our results also imply that stake size may not be a good predictor of monitoring intensity. With proprietary blocks, larger stakes imply more monitoring because stake size directly determines monitoring intensity. However, with delegated blocks the fund's internal incentive structure separates monitoring incentives from stake size. As a result, funds with smaller stakes might actually monitor more intensely than those with larger stakes depending on their investor clientele. Specifically, the total delegated block size in our model is $\omega + \frac{\lambda}{1+\lambda}(1+\omega)$ which is increasing both in the number of FCs and in the aggregate endowment of the FCs. In contrast, the monitoring intensity depends on the skin in the game of the FMs, which in turn depends only on the aggregate endowment of the FCs. Potentially, therefore, funds with many investors holding limited endowments (large λ, small ω) can have large blocks with very little monitoring. In contrast, funds with fewer investors but higher endowments (small λ, relatively large ω) can hold relatively small blocks but monitor significantly more.

As an illustrative example, assuming quadratic monitoring costs $\frac{1}{2}\gamma m^2$, we compare a fund with a relatively large number ($\lambda = 15\%$) of investors with very limited endowments of the risky asset ($\omega = 0.1\%$) to a fund with a relatively small number ($\lambda = 5\%$) of investors with relatively high endowments of the risky asset ($\omega = 0.5\%$). The former fund would hold approximately 13% of the firm, FMs would own a very small fraction—0.8%—of the fund's assets under management, and for $\gamma = 0.1$, monitoring would occur at a low intensity of $\frac{\omega}{\gamma} = 0.01$. The latter fund would hold approximately 5% of the firm, i.e., a much smaller stake; FMs would own a much larger fraction—9.5%—of the fund's assets under management, and for $\gamma = 0.1$, monitoring would occur at five-times the intensity of the other fund, $\frac{\omega}{\gamma} = 0.05$.

While our model is not ideally suited for calibration, it is noteworthy that these block sizes, FM-ownership stakes, and monitoring intensity are broadly in line with observations about mutual fund families and hedge funds. Large families like Blackrock or Fidelity often own well over 10% of US corporations but arguably do not monitor much, while their managers typically have very small stakes in the fund (as discussed above). In contrast, activist hedge funds hold a median stake of around 5-6% in target firms, monitor intensively, and—as discussed above—their general partners often hold a personal stake of around 10% of the fund's assets under management. Our results are also consistent with Nockher (2022), who finds that smaller blockholders, and particularly those with a larger percentage of their fund invested in a given firm, tend to be more engaged monitors than larger blockholders.

7 Conclusion

Blockholder monitoring is important, but the determinants of long-term block sizes and the resulting implications for the degree of monitoring are not fully understood. The existing theoretical literature devoted to this question focuses on proprietary blockholding, whereas modern markets are dominated by delegated asset managers. We present a simple model of delegated trading and monitoring to examine the economics of concentrated ownership and blockholder monitoring in financial markets dominated by institutional investors.

Our analysis shows that delegation has important consequences for both block sizes and monitoring. In particular, optimal delegation contracts allow for the separation of risk sharing and monitoring motives. This can lead to less monitoring and inferior risk sharing relative to proprietary blocks, but gives rise to monitoring and risk sharing benefits where proprietary blocks would not exist.

At an applied level, our model illustrates how some commonly observed characteristics of asset management firms—the clientele they serve, the extent of managerial self-investment, and the degree to which they monitor portfolio firms—can arise as a result of optimal contracting with fund investors. Further, our results imply that block size may not be a good predictor of monitoring intensity because the fund's internal incentive structure separates monitoring incentives from stake size. Finally, given that we conclude that active asset managers may endogenously avoid utilizing their full risk-bearing capacity to hold concentrated positions, our analysis indirectly highlights the importance of the governance role of index asset managers—who mechanically hold concentrated stakes—in corporate governance.

Appendix

Proof of Proposition 1: We begin with condition (i) of the globally stable allocation. Combining definition (1) with the monitoring level $m(\alpha)$ defined in (2) and the market clearing price (3), the optimization problem can be written as:

$$\max_{\alpha} \alpha\mu(m(\alpha)) - c(m(\alpha)) - \frac{1}{2\lambda}\alpha^2\sigma^2 - \Psi(\alpha_G) - (\alpha - \alpha_G)\left(\mu(m(\alpha_G)) - \frac{1-\alpha_G}{\rho(1-\lambda)}\sigma^2\right),$$

giving rise to the following first order condition:

$$\mu(m(\alpha)) + \alpha\mu'(m(\alpha))m'(\alpha) - c'(m(\alpha))m'(\alpha) - \frac{1}{\rho\lambda}\alpha\sigma^2 - \left(\mu(m(\alpha_G)) - \frac{1-\alpha_G}{\rho(1-\lambda)}\sigma^2\right) = 0.$$

Since $m(\alpha)$ satisfies $\alpha m'(\alpha) - c'(\alpha) = 0$, this simplifies to

$$\mu(m(\alpha)) - \frac{1}{\rho\lambda}\alpha\sigma^2 - \left(\mu(m(\alpha_G)) - \frac{1-\alpha_G}{\rho(1-\lambda)}\sigma^2\right) = 0$$

Now, setting $\alpha = \alpha_G$ above and solving gives:

$$\frac{1}{\rho\lambda}\alpha_G\sigma^2 = \frac{1-\alpha_G}{\rho(1-\lambda)}\sigma^2, \text{i.e., } \alpha_G = \lambda.$$

Now, we turn to condition (ii) of the globally stable allocation to verify that $\Psi(\lambda) - \Psi(\omega) - (\lambda - \omega)P(\lambda) > 0$ for all $\omega \neq \lambda$. This is equivalent to showing that $\omega = \lambda$ is a global maximum of the function

$$\Psi(\omega) - \omega P(\lambda) = \omega\mu(m(\omega)) - c(m(\omega)) - \frac{1}{2\rho\lambda}\omega^2\sigma^2 - \omega\left(\mu(m(\lambda)) - \frac{\sigma^2}{\rho}\right).$$

To verify this we first note that the simplified first order condition

$$\mu(m(\omega)) - \frac{1}{\rho\lambda}\omega\sigma^2 - \left(\mu(m(\lambda)) - \frac{\sigma^2}{\rho}\right) = 0$$

is satisfied at $\omega = \lambda$. We then evaluate the second order condition at $\omega = \lambda$: $\mu'(m(\lambda))m'(\lambda) - \frac{\sigma^2}{\rho\lambda}$. This is strictly negative as long as $\Psi(\cdot)$ is strictly concave as required.$\blacksquare$

Proof of Proposition 2: In analyzing the full-commitment case, we assume that FCs collectively commit to a level of monitoring m which is publicly observed. Further, they also commit publicly to a single round of trade. Now, if they trade to a holding of α, then they will face a price of $\mu(m) - \frac{1-\alpha}{\rho(1-\lambda)}\sigma^2$, generating a payoff of

$$\alpha\mu(m) - c(m) - \frac{1}{2\rho\lambda}\alpha^2\sigma^2 - (\alpha - \omega)\left(\mu(m) - \frac{1-\alpha}{\rho(1-\lambda)}\sigma^2\right).$$

Taking the partial derivative of the objective function with respect to m yields a FOC of

$$\omega\mu'(m) - c'(m) = 0,$$

which does not depend on α. The SOC is clearly satisfied given our assumptions, so the solution is given implicitly by $\omega = \frac{c'(m^C)}{\mu'(m^C)}$.

Taking the partial derivative of the objective function with respect to α for a given m and simplifying yields the FOC

$$\frac{(1+\omega)\sigma^2}{\rho(1-\lambda)} - \alpha\left(\frac{\sigma^2}{\rho\lambda} + \frac{2\sigma^2}{\rho(1-\lambda)}\right) = 0,$$

and again the SOC is clearly satisfied. Solving for α yields an optimal stake of

$$\alpha^C = \frac{1+\omega}{1+\lambda}\lambda. \qquad \blacksquare$$

Proof of Lemma 1: In the full-commitment optimum, each individual nonfinancial investor has a payoff of

$$\frac{1}{\lambda}\left(\alpha^C\mu(m^C) - c(m^C) - \frac{1}{2\rho\lambda}(\alpha^C)^2\sigma^2 - (\alpha^C - \omega)\left(\mu(m^C) - \frac{1-\alpha^C}{\rho(1-\lambda)}\sigma^2\right)\right).$$

If a single nonfinancial investor were to stay out of the fund and instead consume their endowment, they would receive a payoff of

$$\frac{1}{\lambda}\left(\omega\mu(m^C) - \frac{1}{2\rho\lambda}\omega^2\sigma^2\right).$$

Subtracting the latter from the former yields a difference of

$$\frac{1}{\lambda}\left((\alpha^C - \omega)\mu(m^C) - ((\alpha^C)^2 - \omega^2)\frac{\sigma^2}{2\rho\lambda} - c(m^C) - (\alpha^C - \omega)\left(\mu(m^C) - \frac{1-\alpha^C}{\rho(1-\lambda)}\sigma^2\right)\right),$$

which is clearly negative when $\omega = \lambda$ (in which case $\alpha^C = \lambda$), and clearly positive when $\omega = 0$ (by virtue of the definition of α^C and m^C). Thus, if the difference decreases monotonically in ω, then by continuity there will be exactly one value of $\omega \in (0, \lambda)$ for which the difference is exactly zero. We take the $\omega-$derivative of the difference while accounting for the dependence of α^C and m^C on ω. This yields $\frac{1}{\lambda}$ times

$$\left(\frac{\lambda}{1+\lambda} - 1\right)\mu(m^C) + \left(\frac{1+\omega}{1+\lambda}\lambda - \omega\right)\mu'(m^C)\frac{\partial m^C}{\partial\omega} - \frac{\sigma^2}{2\rho\lambda}\left(2\left(\frac{1+\omega}{1+\lambda}\lambda\right)\left(\frac{\lambda}{1+\lambda}\right) - 2\omega\right) - c'(m^C)\frac{\partial m^C}{\partial\omega}$$

$$-\left(\frac{\lambda}{1+\lambda} - 1\right)\left(\mu(m^C) - \frac{1 - \frac{1+\omega}{1+\lambda}\lambda}{\rho(1-\lambda)}\sigma^2\right) - \left(\frac{1+\omega}{1+\lambda}\lambda - \omega\right)\left(\mu'(m^C)\frac{\partial m^C}{\partial\omega} + \frac{\lambda}{1+\lambda}\frac{\sigma^2}{\rho(1-\lambda)}\right)$$

or

$$-\frac{\sigma^2}{2\rho\lambda}\left(2\left(\frac{1+\omega}{1+\lambda}\lambda\right)\left(\frac{\lambda}{1+\lambda}\right) - 2\omega\right) - c'(m^C)\frac{\partial m^C}{\partial\omega}$$

$$-\left(\frac{\lambda}{1+\lambda} - 1\right)\left(-\frac{1 - \frac{1+\omega}{1+\lambda}\lambda}{\rho(1-\lambda)}\sigma^2\right) - \left(\frac{1+\omega}{1+\lambda}\lambda - \omega\right)\left(\frac{\lambda}{1+\lambda}\frac{\sigma^2}{\rho(1-\lambda)}\right)$$

or

$$-\frac{\sigma^2(\lambda - \omega)}{\lambda(1-\lambda^2)\rho} - c'(m^C)\frac{\partial m^C}{\partial\omega} < 0, \textbf{since } \frac{\partial m^C}{\partial\omega} > 0. \qquad \blacksquare$$

Proof of Lemma 2: We begin with condition (i) of the globally stable allocation. Combining definition (4) with the selected monitoring level $m^D(\alpha^D)$ defined by $\phi\alpha^D = \frac{c'(m^D(\alpha^D))}{\mu'(m^D(\alpha^D))}$ and the market clearing price (5), the optimization problem can be written as:

$$\max_{\alpha^D} \phi\alpha^D\mu(m^D(\phi\alpha^D))-c(m^D(\phi\alpha^D))-\frac{\phi^2(\alpha^D)^2\sigma^2}{2\rho\tau}-\Psi(\alpha_G^D)-\phi(\alpha^D-\alpha_G^D)\left(\mu(m^D(\phi\alpha_G^D))-\frac{1-\alpha_G^D}{\rho(1-\lambda-\tau)}\sigma^2\right),$$

giving rise to the following first order condition:

$$\phi\mu(m^D(\phi\alpha^D))-\frac{1}{\rho\tau}\phi^2\alpha^D\sigma^2-\phi\left(\mu(m^D(\phi\alpha_G^D))-\frac{1-\alpha_G^D}{\rho(1-\lambda-\tau)}\sigma^2\right)=0.$$

Now, setting $\alpha^D=\alpha_G^D$ above and solving gives

$$\frac{1}{\rho\tau}\phi^2\alpha_G^D\sigma^2=\phi\frac{1-\alpha_G^D}{\rho(1-\lambda-\tau)}\sigma^2, \text{i.e., } \alpha_G^D=\frac{\tau/\phi}{\tau/\phi+1-\lambda-\tau}.$$

Now, we turn to condition (ii) of the globally stable allocation to verify that $\Psi^D(\frac{\tau/\phi}{\tau/\phi+1-\lambda-\tau})-\Psi^D(\omega)-\phi(\frac{\tau/\phi}{\tau/\phi+1-\lambda-\tau}-\omega)P^D(\frac{\tau/\phi}{\tau/\phi+1-\lambda-\tau})>0$ for all $\omega\neq\frac{\tau/\phi}{\tau/\phi+1-\lambda-\tau}$. This is equivalent to showing that $\omega=\frac{\tau/\phi}{\tau/\phi+1-\lambda-\tau}$ is a global maximum of the function $\Psi^D(\omega)-\phi\omega P^D(\frac{\tau/\phi}{\tau/\phi+1-\lambda-\tau})$, i.e.,

$$\phi\omega\mu(m^D(\phi\omega))-c(m^D(\phi\omega))-\frac{1}{2\rho\tau}\omega^2\phi^2\sigma^2-\phi\omega\left(\mu\left(m^D(\phi\frac{\tau/\phi}{\tau/\phi+1-\lambda-\tau})\right)-\frac{1}{\rho(\tau/\phi+1-\lambda-\tau)}\sigma^2\right).$$

To verify this we first note that the first order condition

$$\phi\mu(m^D(\phi\omega))-\frac{1}{\rho\tau}\omega\phi^2\sigma^2-\phi\left(\mu\left(m^D(\phi\frac{\tau/\phi}{\tau/\phi+1-\lambda-\tau})\right)-\frac{1}{\rho(\tau/\phi+1-\lambda-\tau)}\sigma^2\right)=0$$

is satisfied at $\omega=\frac{\tau/\phi}{\tau/\phi+1-\lambda-\tau}$. We then evaluate the second order condition at $\omega=\frac{\tau/\phi}{\tau/\phi+1-\lambda-\tau}$: $\phi\mu'(m^D(\phi\frac{\tau/\phi}{\tau/\phi+1-\lambda-\tau}))m^{D'}(\phi\frac{\tau/\phi}{\tau/\phi+1-\lambda-\tau})-\frac{\phi^2\sigma^2}{\rho\tau}$. This is strictly negative as long as $\Psi^D(\cdot)$ is strictly concave as required.■

Proof of Lemma 3: The per-FM effective allocation is $\frac{\phi\alpha_G^D}{\tau}=\frac{\phi}{\tau+\phi-\phi(\lambda+\tau)}$. The financial investors outside the fund hold an aggregate stake of $1-\alpha_G^D=\frac{\phi-\phi(\lambda+\tau)}{\tau+\phi-\phi(\lambda+\tau)}$, leading to a per-investor allocation of $\frac{1}{1-\lambda-\tau}\frac{\phi-\phi(\lambda+\tau)}{\tau+\phi-\phi(\lambda+\tau)}=\frac{\phi}{\tau+\phi-\phi(\lambda+\tau)}$. ■

Proof of Proposition 3: To replicate a payoff of Π_{FC}^{C} for the FCs, (1) the fund, i.e., the FMs, must choose to monitor at level m^C, which they will only do if their own stake inside the fund is equal to ω units of the risky asset; and (2) the FCs must hold a final stake inside the fund of $\alpha^C = \frac{\lambda(1+\omega)}{(1+\lambda)}$ units of the risky asset. We choose ϕ and τ to achieve (1) and (2). For (1), we require that $\phi\alpha_G^D = \omega$. For (2), we require that $(1-\phi)\alpha_G^D = \frac{\lambda(1+\omega)}{(1+\lambda)}$. Plugging in the definition of α_G^D and solving these two equations for the two unknowns ϕ and τ yields ϕ^* and τ^* as given in the text of Proposition 3. From here forward, let the superscript $D*$ indicate that the associate function or variable is evaluated at τ^* and ϕ^*.

Next we determine the fee level f^* that just meets the participation constraint of individual FMs to ensure that the optimal mass τ^* will join the fund given the optimal skin in the game parameter ϕ^*. The fund's total endowment is given by $\omega + \tau^* \frac{1-\omega}{1-\lambda}$ (the FCs' endowment plus the FMs' share of the financial investors' aggregate endowment of $1-\omega$). The per-FM payoff for those who join the fund is given by

$$\frac{1}{\tau^*}\left[\Psi^{D*}(\alpha_G^{D*}) - \phi^*\left(\alpha_G^{D*} - \omega - \tau^*\frac{(1-\omega)}{1-\lambda}\right)P^{D*}(\alpha_G^{D*}) + \lambda f\right]. \tag{8}$$

The per-investor payoff for financial investors who do not join the fund is

$$\frac{1}{1-\lambda-\tau^*}\left[\Psi_U^{D*}(\alpha_G^{D*}) - \left(1-\alpha_G^{D*} - \left(1-\omega-\tau^*\frac{(1-\omega)}{1-\lambda}\right)\right)P^{D*}(\alpha_G^{D*})\right],$$

where $\Psi_U^D(\alpha^D) = (1-\alpha^D)\mu(m^D(\alpha^D)) - \frac{(1-\alpha^D)^2\sigma^2}{2\rho(1-\lambda-\tau)}$ is the aggregate certainty equivalent payoff of the mass of $1-\lambda-\tau$ financial investors outside of the fund who hold an aggregate stake of $1-\alpha^D$ given that the fund holds a stake of α^D. By defecting from the fund unilaterally, any given FM who is supposed to join the fund can realize the latter payoff. Thus, their participation constraint will be met as long as f is set to make these two payoffs equivalent.

Defining

$$f^* = \frac{1}{\lambda}\left(\begin{array}{c} \frac{\tau^*}{1-\lambda-\tau^*}\left[\Psi_U^{D*}(\alpha_G^{D*}) - \left(1-\alpha_G^{D*} - \left(1-\omega-\tau^*\frac{(1-\omega)}{1-\lambda}\right)\right)P^{D*}(\alpha_G^{D*})\right] \\ -\left[\Psi^{D*}(\alpha_G^{D*}) - \phi^*\left(\alpha_G^{D*}-\omega-\tau^*\frac{(1-\omega)}{1-\lambda}\right))P^{D*}(\alpha_G^{D*})\right] \end{array} \right) \tag{9}$$

ensures the participation of the requisite mass of FMs. Later in this proof, we show that the above expression for $f^*$is equivalent to the expression shown in the statement of Proposition 3.

To complete the proof, we now show that the above contracting terms lead to an aggregate payoff for the FCs of Π_{FC}^C. First, we show that the price of the risky asset in the delegated fund equilibrium is equivalent to the price in the FCs' full-commitment optimum. In the full-commitment optimum, the price is given by $\mu(m^C) - \frac{1-\alpha^C}{\rho(1-\lambda)}\sigma^2$. Replacing α^C with $\frac{\lambda(1+\omega)}{(1+\lambda)}$ yields a price of $\mu(m^C) - \frac{1-\lambda\omega}{\rho(1-\lambda^2)}\sigma^2$. In the delegated fund equilibrium, the price, evaluated at the optimal fund parameters, is given by $P^{D*}(\alpha_G^{D*}) = \mu(m^C) - \frac{1-\frac{\tau^*/\phi^*}{\tau^*/\phi^*+1-\lambda-\tau^*}}{\rho(1-\lambda-\tau^*)}\sigma^2$. It is straightforward to show the algebraic equivalence of these two prices using the definitions of τ^* and ϕ^*.

Since the level of monitoring in the delegated fund equilibrium is identical to that in the FCs' full-commitment optimum and the FCs' final holdings are identical, to complete our argument it suffices to show that FCs pay identical effective monitoring costs and trading costs across the two cases. With respect to the monitoring costs, note that the FMs directly pay the entirety of the actual costs in the delegated fund equilibrium while the FCs pay these costs in the full-commitment optimum. Thus, the aggregate fee paid by the FCs must compensate FMs for their monitoring costs. With respect to trading costs, the costs incurred by the FCs in their full-commitment optimum equal the equilibrium price times their aggregate trading quantity, or $P^{D*}(\alpha_G^{D*})(\alpha^C - \omega)$ (using the result above that the equilibrium price is equivalent to the full-commitment price). In the delegated fund equilibrium they directly pay trading costs equal to the price times their proportional stake in the fund times its overall trading quantity, or $P^{D*}(\alpha_G^{D*})(1-\phi^*)(\alpha_G^{D*} - \omega - \tau^*\frac{(1-\omega)}{1-\lambda})$. Since, as shown previously, $(1-\phi^*)\alpha_G^{D*} = \alpha^C$, the *savings* in the FCs' direct trading costs for the delegated fund equilibrium relative to their full-commitment equilibrium equal

$$P^{D*}(\alpha_G^{D*})\left[(\alpha^C-\omega)-(1-\phi^*)(\alpha_G^{D*}-\omega-\tau^*\frac{(1-\omega)}{1-\lambda})\right]=P^{D*}(\alpha_G^{D*})\left[(1-\phi^*)(\omega+\tau^*\frac{(1-\omega)}{1-\lambda})-\omega\right].$$

Thus, the aggregate fee must also transfer this amount from the FCs to the FMs.

To show that the equilibrium fee, f^* as defined in (9) accomplishes these requirements, first note that it is straightforward to show that $\frac{\tau^*}{1-\lambda-\tau^*}(1-\alpha_G^{D*})=\omega$, and since we know that $\phi^*\alpha_G^{D*}=\omega$ also holds, we have $\frac{\tau^*}{1-\lambda-\tau^*}\Psi_U^{D*}(\alpha_G^{D*})-\Psi^{D*}(\alpha_G^{D*})=c(m^C)$. We can therefore rewrite the fee f^* as

$$\frac{1}{\lambda}\left[c(m^C)+P^{D*}(\alpha_G^{D*})\left(\phi^*(\alpha_G^{D*}-\omega-\tau^*\frac{(1-\omega)}{1-\lambda})-\frac{\tau^*}{1-\lambda-\tau^*}\left(1-\alpha_G^{D*}-\left(1-\omega-\tau^*\frac{(1-\omega)}{1-\lambda}\right)\right)\right)\right]$$

or

$$\frac{1}{\lambda}\left[c(m^C)+P^{D*}(\alpha_G^{D*})\left(\phi^*(-\omega-\tau^*\frac{(1-\omega)}{1-\lambda})-\frac{\tau^*}{1-\lambda-\tau^*}\left(-\left(1-\omega-\tau^*\frac{(1-\omega)}{1-\lambda}\right)\right)\right)\right]$$

or

$$\frac{1}{\lambda}\left[c(m^C)+P^{D*}(\alpha_G^{D*})\left(\phi^*(-\omega-\tau^*\frac{(1-\omega)}{1-\lambda})-\frac{\phi^*\alpha_G^{D*}}{(1-\alpha_G^{D*})}\left(-\left(1-\omega-\tau^*\frac{(1-\omega)}{1-\lambda}\right)\right)\right)\right]$$

or

$$\frac{1}{\lambda}\left[c(m^C)+P^{D*}(\alpha_G^{D*})\left((1-\phi^*)(\omega+\tau^*\frac{(1-\omega)}{1-\lambda})-\omega\right)\right],$$

as in Proposition 3. Thus, since the aggregate fee is λf^*, the FCs' obtain payoff Π_{FC}^C. Finally, note that Lemma 1 implies that all λ FCs find participation in the fund optimal as long as $\omega\leq\hat{\omega}$.■

Proof of Corollary 1: The fund holds $\alpha_G^{D*}=\frac{\tau^*/\phi^*}{\tau^*/\phi^*+1-\lambda-\tau^*}$ of the risky asset in equilibrium. The fund is made up of investors of measure $\lambda+\tau^*$ and thus the collective risk tolerance of this group of investors is $\lambda+\tau^*$. In a competitive equilibrium, such measure of investors will hold $\lambda + \tau^*$ of the risky asset. Using the expressions in Proposition 3 we have:

$$\frac{\tau^*}{\phi^*} = \frac{\left(1-\lambda^2\right)\omega}{1-\lambda\omega}\frac{2\lambda\omega+\lambda+\omega}{(1+\lambda)\omega} = \frac{(1-\lambda)\left(2\lambda\omega+\lambda+\omega\right)}{1-\lambda\omega}.$$

We first show that $\tau^*/\phi^* < \lambda + \tau^*$. Assume the contrary. This implies that:

$$\frac{(1-\lambda)\left(2\lambda\omega+\lambda+\omega\right)}{1-\lambda\omega} \geq \lambda + \frac{\left(1-\lambda^2\right)\omega}{1-\lambda\omega},$$

which simplifies to $\lambda\left(\omega-\lambda\right)\geq 0$, which is a contradiction because $\lambda > 0$ and $\omega \leq \lambda$. Having shown that $\frac{\tau^*}{\phi^*} < \lambda+\tau^*$, we now observe that $\alpha_G^{D*} = \frac{\tau^*/\phi^*}{\tau^*/\phi^*+1-\lambda-\tau^*} < \frac{\lambda+\tau^*}{\lambda+\tau^*+1-\lambda-\tau^*} = \lambda+\tau^*$.■

Proof of Lemma 4: From above, the payoff of an FC in the fund is given by

$$\frac{1}{\lambda}\left(\alpha^C\mu(m^C) - c(m^C) - \frac{1}{2\rho\lambda}(\alpha^C)^2\sigma^2 - (\alpha^C-\omega)\left(\mu(m^C) - \frac{1-\alpha^C}{\rho(1-\lambda)}\sigma^2\right)\right),$$

whereas the payoff of a financial investor outside the fund is

$$\frac{1}{1-\lambda-\tau^*}\left[\Psi_U^{D*}(\alpha_G^{D*}) - \left(1-\alpha_G^{D*} - \left(1-\omega-\tau^*\frac{(1-\omega)}{1-\lambda}\right)\right)P^{D*}(\alpha_G^{D*})\right].$$

A nonfinancial investor who pays the opportunity cost, however, starts from a smaller endowment, so its payoff outside the fund is

$$\frac{1}{1-\lambda-\tau^*}\left[\Psi_U^{D*}(\alpha_G^{D*}) - \left(1-\alpha_G^{D*} - \frac{\omega}{\lambda}(1-\lambda-\tau^*)\right)P^{D*}(\alpha_G^{D*})\right].$$

Replacing $\Psi_U^{D*}(\alpha_G^{D*})$ with its definition $(1-\alpha_G^{D*})\mu(m^D(\alpha_G^{D*})) - \frac{(1-\alpha^{D*})^2\sigma^2}{2\rho(1-\lambda-\tau)}$, recognizing that $\alpha_G^{D*} = \alpha^C+\omega$ and $P^{D*}(\alpha_G^{D*}) = \mu(m^C) - \frac{1-\alpha^C}{\rho(1-\lambda)}\sigma^2$, gives

$$\frac{1}{1-\lambda-\tau^*}\left[\begin{array}{c} (1-(\alpha^C+\omega))\mu(m^C) - \frac{(1-(\alpha^C+\omega))^2\sigma^2}{2\rho(1-\lambda-\tau^*)} \\ -\left(1-(\alpha^C+\omega)-\frac{\omega}{\lambda}(1-\lambda-\tau^*)\right)\left(\mu(m^C)-\frac{1-\alpha^C}{\rho(1-\lambda)}\sigma^2\right)\end{array}\right].$$

Replacing τ^* with $\frac{(1-\lambda^2)\omega}{1-\lambda\omega}$ and α^C with $\frac{\lambda(1+\omega)}{(1+\lambda)}$ and simplifying gives

$$\frac{1-\lambda\omega}{1-\lambda^2}\left(\mu(m^C)-\left(\frac{1-\lambda\omega}{1-\lambda^2}\right)\frac{\sigma^2}{2\rho}\right)-\left(\frac{1-\lambda\omega}{1-\lambda^2}-\frac{\omega}{\lambda}\right)\left(\mu(m^C)-\frac{1-\frac{\lambda(1+\omega)}{(1+\lambda)}}{\rho(1-\lambda)}\sigma^2\right).$$

Replacing α^C with $\frac{\lambda(1+\omega)}{(1+\lambda)}$ in the per-FC payoff expression above and simplifying yields

$$\frac{1+\omega}{1+\lambda}\left(\mu(m^C)-\left(\frac{1+\omega}{1+\lambda}\right)\frac{\sigma^2}{2\rho}\right)-\left(\frac{(1+\omega)}{(1+\lambda)}-\frac{\omega}{\lambda}\right)\left(\mu(m^C)-\frac{1-\frac{\lambda(1+\omega)}{(1+\lambda)}}{\rho(1-\lambda)}\sigma^2\right)-\frac{c(m^C)}{\lambda}.$$

Subtracting the latter from the former (noting that $\frac{1-\lambda\omega}{1-\lambda^2}-\frac{1+\omega}{1+\lambda}=\frac{\lambda-\omega}{1-\lambda^2}>0$) yields

$$\left(\frac{\lambda-\omega}{1-\lambda^2}\right)\left(\frac{1-\lambda\omega}{\rho(1-\lambda^2)}\sigma^2\right)-\left(\frac{\lambda-\omega}{1-\lambda^2}\right)\left(\frac{2-\lambda+\omega-2\lambda\omega}{1-\lambda^2}\right)\frac{\sigma^2}{2\rho}+\frac{c(m^C)}{\lambda},$$

or, simplifying,

$$\left(\frac{\lambda-\omega}{1-\lambda^2}\right)^2\frac{\sigma^2}{2\rho}+\frac{c(m^C)}{\lambda},$$

as required.■

Proof of Proposition 4: Consider any arbitrary starting point where an optimal contract is in place under which monitoring occurs at intensity corresponding to a block size of $\omega<\hat{\omega}<\lambda$ and FCs holdings correspond to $\frac{\lambda}{1+\lambda}(1+\omega)$. Consider an ex post recontracting opportunity. There are two relevant cases:

1. $\frac{\lambda}{1+\lambda}(1+\omega)\leq\hat{\omega}$ or
2. $\frac{\lambda}{1+\lambda}(1+\omega)>\hat{\omega}$.

In case (1), a new optimal fund can be formed, where monitoring occurs at a higher level corresponding to a block size of $\frac{\lambda}{1+\lambda}(1+\omega)$, and FC holdings correspond to $\frac{\lambda}{1+\lambda}(1+\frac{\lambda}{1+\lambda}(1+\omega))$. However, since here $\frac{\lambda}{1+\lambda}(1+\omega)\leq\hat{\omega}<\lambda$, monitoring is clearly lower than the APZ outcome, and since $\frac{\lambda}{1+\lambda}(1+x)$ is increasing in x, and $\frac{\lambda}{1+\lambda}(1+\omega)\leq\hat{\omega}$, thus $\frac{\lambda}{1+\lambda}(1+\frac{\lambda}{1+\lambda}(1+\omega))<\frac{\lambda}{1+\lambda}(1+\hat{\omega})\leq\lambda$, and thus the risk sharing achieved by FCs is also lower than in the APZ outcome. The new optimal fund then again can belong to case (1) (in which case the previous argument applies) or belong to case (2) in which the argument below will apply.

In case (2), FC holdings are so high that by Lemma 1 it is not possible to form the optimal fund. Can some potentially suboptimal fund be formed that takes us closer to the APZ outcome? Since $\omega < \lambda$, monitoring is lower than in the APZ outcome. Hence to reach the APZ outcome, let us consider enhancing monitoring to some arbitrary level corresponding to a holding of $\omega' > \omega$. Denote by $m(\omega')$ the monitoring achieved by the proposed fund. At this level of monitoring, the proposed suboptimal fund must offer sufficient risk sharing for FCs not to leave the fund in order to enjoy its monitoring for free. What is the maximal risk sharing benefit that can be offered by any fund to FCs? By reference to Proposition 2, at an endowment level of $\frac{\lambda}{1+\lambda}(1+\omega)$, the maximal risk sharing benefit, net of trading costs, that can be achieved at a monitoring level corresponding to holding level ω' is given by the maximal value of :

$$max_{\alpha}\alpha\mu(m(\omega')) - c(m(\omega')) - \frac{1}{2\rho\lambda}\alpha^2\sigma^2 - (\alpha - \frac{\lambda}{1+\lambda}(1+\omega))\left(\mu(m(\omega')) - \frac{1-\alpha}{\rho(1-\lambda)}\sigma^2\right),$$

which is attained at

$$\alpha^C = \frac{\lambda}{1+\lambda}(1+\frac{\lambda}{1+\lambda}(1+\omega)).$$

But that is simply the optimal fund's risk sharing allocation. Yet, since $\frac{\lambda}{1+\lambda}(1+\omega) > \hat{\omega}$, Lemma 1 shows that such a risk sharing benefit will not induce FCs to join the fund even at a lower monitoring level, corresponding to a holding of $\omega < \omega'$. Clearly, then, FCs will not agree to join the fund with even higher monitoring, and no new fund can be formed.■

Proof of Proposition 5: Suppose a fund involving $\theta \in (0, 1-\lambda]$ financial investors is proposed and would hold a total (post-trade) stake of α_θ (held uniformly across the financial investors inside the fund) and monitor according to their collective holdings. Each financial investor must pay some part of (i) the cost of trading to α_θ and (ii) the total monitoring cost of $c(m(\alpha_\theta))$where $m(\alpha_\theta)$ is implicitly defined by the solution to: $\alpha_\theta = \frac{c'(m(\alpha_\theta))}{\mu'(m(\alpha_\theta))}$. Instead, any financial investor could leave the fund, enjoy the fund's monitoring for free (thereby saving on (ii)), and trade to a holding of $\frac{\alpha_\theta}{\theta}$ at a $\frac{1}{\theta}$-fraction of the total cost in (i). Any financial investor who was paying more than the average trading cost would then benefit from not joining the fund. Thus, all financial investors must share in the trading cost equally. Given this, each financial investor who shared in the monitoring cost would find it strictly beneficial to not join the fund, as their trading cost would be identical to their cost inside the fund but they would save on monitoring. The only financial investors who would not strictly prefer to stay out of the fund would be ones who paid no monitoring costs. But if nobody paid monitoring costs, then the fund could not monitor according to the collective holdings of financial investors inside the fund, which is a contradiction.■

Proof of Proposition 6. Consider a fund with a measure λ FCs and measure τ FMs, with the latter owning a fraction ϕ of the fund. Let $\alpha_G^D(\tau,\phi)$ denote the risky asset holdings of the fund, given by the globally stable allocation (6), and let $\alpha^{lim}(0,0) := lim_{\tau\to 0,\phi\to 0}\alpha_G^D(\tau,\phi)$ (if it exists) denote the risky asset holdings of the fund in the limit as $\tau \to 0$ and $\phi \to 0$. In such a limiting fund, FCs own everything (because $\phi \to 0$) and there is no monitoring (because $\tau \to 0$). Thus, by joining the fund the FCs obtain a payoff of

$$CE_{FC}^{Fund} := \alpha^{lim}(0,0)\mu(0) - \frac{1}{2\rho\lambda}(\alpha^{lim}(0,0))^2\sigma^2 - (\alpha^{lim}(0,0)-\omega)(\mu(0) - \frac{1-\alpha^{lim}(0,0)}{\rho(1-\lambda)}\sigma^2),$$

whereas if they do not join the fund they enjoy a payoff of

$$\omega\mu(0) - \frac{1}{2\rho\lambda}\omega^2\sigma^2.$$

Thus, the FC's willingness to pay for joining this fund is:

$$WTP_{FC}(\alpha^{lim}(0,0)) = CE_{FC}^{Fund} - \omega\mu(0) + \frac{1}{2\rho\lambda}\omega^2\sigma^2.$$

Clearly, $WTP_{FC}(\alpha^{lim}(0,0))$ is maximized at $\alpha^{lim}(0,0) = \frac{1+\omega}{1+\lambda}\lambda$ (the derivation can be found in the proof of Proposition 2).

The monopolistic fund designer can only extract risk-sharing rents from FCs. This is because any monitoring provided by the fund can be enjoyed by FCs outside the fund. Any monitoring undertaiken by the fund must therefore be paid for by the fund designer out of fund fees in order to satisfy the FMs participation constraint, because, unless FMs were compensated for their monitoring costs they would refuse to be fund managers. As a result, the monopolistic fund designer can optimize rents by maximizing the risk sharing attained by FCs in a fund that undertakes no monitoring, i.e., by designing a fund, if feasible, in which $\tau \to 0$, $\phi \to 0$, and $lim_{\tau \to 0, \phi \to 0} \alpha_G^D(\tau, \phi) = \frac{1+\omega}{1+\lambda}\lambda$.

To demonstrate feasibility, for any τ,let $\phi = \frac{1-\omega\lambda}{\lambda(1-\lambda)(1+\omega)}\tau$. Then, for each $\tau > 0$, we have

$$\alpha_G^D = \frac{\frac{\lambda(1-\lambda)(1+\omega)}{1-\omega\lambda}}{\frac{\lambda(1-\lambda)(1+\omega)}{1-\omega\lambda} + 1 - \lambda - \tau}.$$

Now, taking the limit as $\tau \to 0$ and simplifying gives $lim_{\tau \to 0, \phi \to 0} \alpha_G^D(\tau, \phi) = \frac{1+\omega}{1+\lambda}\lambda$ as required.■

Bibliography

Admati, A., P. Pfleiderer, and J. Zechner 1994. Large shareholder activism, risk sharing, and financial market equilibrium. Journal of Political Economy 102, 1097-1130.

Agarwal, Vikas, Naveen D. Daniel, and Narayan Y. Naik, 2009, Role of managerial incentives and discretion in hedge fund performance, Journal of Finance 64, 2221–2256.

Bebchuk, L. A., A. Cohen, and S. Hirst 2017. The agency problems of institutional investors. Journal of Economic Perspectives. 31, 89–102.

Bhattacharya, S. and P. Pfleiderer 1985. Delegated Portfolio Management. Journal of Economic Theory 36, 1-25.

Bolton, P. and H. Rosenthal 2019. The End of the "Modern Corporation": Deregulation and Ownership of Electric Utilities. Working paper, Columbia University.

Bolton, P. and E. von Thadden 1998. Blocks, Liquidity, and Corporate Control. Journal of Finance 53, 1-25.

Brav, A., W. Jiang, and H. Kim 2010. Hedge fund activism: A review. Foundations and Trends in Finance. 4, 185–246.

Brav, A., A. Malenko, and N. Malenko 2022. Corporate governance implications of the growth in indexing. ECGI working paper 849/2022.

Corum, A., A. Malenko, and N. Malenko 2023. Corporate governance in the presence of active and passive delegated investment. ECGI working paper 695/2020.

Dasgupta A., V. Fos, Z. Sautner 2021. Institutional Investors and Corporate Governance. Foundations and Trends in Finance, 12, 276–394.

Dasgupta, A., and E. Maug 2023. Delegation chains. ECGI working paper 798/2021.

Dasgupta, A. and G. Piacentino 2015. The wall street walk when blockholders compete for flows. Journal of Finance 70, 2853– 2896.

DeMarzo, P. and B. Urosevic 2006. Ownership Dynamics and Asset Pricing with a Large Shareholder. Journal of Political Economy 114, 774-815.

Diamond, D. and P. Dybvig 1983. Bank Runs, Deposit Insurance, and Liquidity. Journal of Political Economy 91, 401–419.

Grossman, Sanford J., and Oliver D. Hart, 1980, Takeover bids, the free-rider problem, and the theory of the corporation, Bell Journal of Economics 11, 42–64.

He, Zhiguo, and Arvind Krishnamurthy, 2013, Intermediary asset pricing. American Economic Review 103, 732–770.

Khorana, Ajay, Henri Servaes, and Lei Wedge, 2007, Portfolio manager ownership and fund performance. Journal of Financial Economics 85, 179–204.

Kyle, A. and J.-L.Vila 1991. Noise trading and takeovers. RAND Journal of Economics. 22, 54–71.

Marinovic, I. and F. Varas 2021. Strategic trading and blockholder dynamics. Working paper, Stanford University.

Nockher, F. 2022. The Value of Undiversified Shareholder Engagement. Working paper, University of Pennsylvania.

Piacentino, G. 2019, Venture capital and capital allocation. Journal of Finance 74, 1261-1314.

Shleifer. A. and R. Vishny 1986. Large shareholders and corporate control. Journal of Political Economy 94, 461–488.