

Mitigating Moral Hazard in Delegated Investment through Recommendation Algorithms ^{*}

Kai Feng[†]

Zhiheng He[†]

Wenshi Wei[†]

July 26, 2025

Digital platforms are increasingly serving as intermediaries in delegated investment, particularly by adopting recommendation algorithms that deliver personalized suggestions to a large user base. We develop a model to analyze the platform’s investor-optimal algorithm design where investors with heterogeneous risk aversion contract with a portfolio manager based on recommendation status. Investors may have limited knowledge about their types, while the manager has risk-chasing incentives due to limited liability. We demonstrate that algorithms can mitigate managers’ moral hazard in over risk-taking — without affecting the contract — by acting as both information gatekeepers and commitment devices, harnessing the scale of the user base. Optimal recommendation probabilities are non-monotonic in historical returns. Unlike in consumption platforms, algorithms here extract noisy signals about the manager’s actions from historical returns, reduce recommendations under ambiguous signals, and potentially compensate for clear signals, leading to an information rent paid by investors. We further discuss on algorithmic inequality, the joint design of algorithms and contracts, and comparisons to fund ranking systems. Our results emphasize the innovative role of recommendation algorithms as a digital financial service. Methodologically, we provide a general approach for algorithm design problems in function space with potentially non-monotonic solutions.

Keywords: FinTech platform, recommendation algorithm, delegated investment, moral hazard, risk chasing

^{*}We are grateful to Bo Hu, Zanhui Liu, Dan Luo, Alexander Rodivilov, Tat-How Teh, Yifan Tao, Stéphane Vileneuve, Longtian Zhang, and participants at conferences and workshops at the 17th Digital Economics Conference at Toulouse School of Economics, 2024 European Winter Meeting of the Econometric Society, 2024 China Tech-Fin Research Conference, 2024 China Information Economics Society Annual Meeting, and HKUST(GZ) for helpful comments. Zhiheng He thanks for the funding support of National Natural Science Foundation of China (Grant No. 723B2013). fengk22@mails.tsinghua.edu.cn (Kai Feng), hezh19@mails.tsinghua.edu.cn (Zhiheng He), wws22@mails.tsinghua.edu.cn (Wenshi Wei).

[†]Institute of Economics, Tsinghua University, 30 Shuangqing Rd., Beijing, 100084, China.

1 Introduction

We are increasingly coming to understand that financial technologies, which have been widely used in asset management, can reshape market participants and information structures remarkably. See, for example [Cookson et al. \(2021\)](#); [Capponi et al. \(2022\)](#); [Hong et al. \(2024\)](#). Platforms collect investment opportunities and provide advisory services, thereby connecting numerous investors with portfolio managers (e.g., mutual fund managers, and active investment agents), aggregating a vast market scale. In the US, prominent examples include traditional institutions, such as Bank of America’s Merrill Guided Investing, Wells Fargo’s Intuitive Investor, as well as fintech entrants like Yieldstreet; In the UK, about 50% of retail mutual fund flows are channeled through investment platforms ([Cookson et al., 2021](#)); in China, Ant Group has covered almost the entire mutual fund market.¹ However, the fact that investors’ attention is concentrated on these platforms creates strong incentives for managers to increase their visibility. With limited liability in delegated investment, managers are motivated to be risk-chasing in pursuit of standout performance ([Hong et al., 2024](#)). The combination of expanded financial inclusion and managerial conflicts of interest can be dangerous, as it leads to numerous non-professional investors becoming overexposed to market risk.

This paper analyzes whether and how these platforms can address the above challenge by leveraging emerging digital technologies. Specifically, we demonstrate that a recommendation algorithm intermediary can mitigate the moral hazard of managers’ excessive risk-taking. This algorithm is predetermined and publicly known. It sends each investor a personalized recommendation signal for a manager (portfolio) based on investor characteristics and the noisy historical portfolio performance. In this way, the algorithm shapes the information structure of the delegated investment without altering the contractual relationship between investors and managers. This offers a potential solution to the agency problem. While this algorithm is designed to improve user-side welfare, our analyzing framework is inherently general and adapted to analyze platform algorithm design with manager-side incentives.²

In our baseline analysis, we incorporate three practically prevalent frictions. (i) Investor heterogeneity and imperfect self-awareness: investors differ in their risk aversion and may have limited knowledge about it ([Capponi et al., 2022](#)). (ii) Fixed contractual structure: the platform cannot modify the contracts between investors and managers. (iii) Opaque manager actions: managers can always (partially) hide their allocation information, even in the presence of disclosure policies,

¹See Appendix C for more institutional background.

²The emerging fintech platforms typically earn revenue from user fees, such as Ant Financial. Consequently, these platforms are incentivized to curb fund managers’ risk-taking and protect user welfare. However, in practice, platforms may also generate income from fund managers by influencing investment flows ([Berk and Green, 2004](#)). In this way, their incentive structure shifts: they must balance user protection with maintaining cooperation with fund providers.

whereas historical performance only provides a noisy signal about the allocation choice.³ We also explore their potential relaxations in extended discussions.

Consider a two-period model where a continuum of non-professional investors contract with a fund manager via a platform. Heterogeneous investors know the distribution of their risk preferences but not their specific risk aversion level. The risk-neutral manager designs the portfolio to maximize the expected total delegation earnings depending on a contract. The fraction of risky asset is private information. The platform uses data to observe investors' risk aversion levels and design a public known algorithm. The algorithm provides a recommendation probability based on an individual's risk aversion and the historical performance of the portfolio.

Intuitively, designing the recommendation algorithm effectively designs the information structure. Without an algorithm, a monotonic and limited-liability contract incentivize the risk-neutral manager to invest all funds in risky assets. However, with an algorithm, the manager knows that the realized return provides a noisy signal about the allocation choice, and the algorithm may penalize suspected excessive risk-taking by reducing recommendation probabilities.

In this paper, we show that a non-monotonic algorithm breaks the monotonicity of the manager's expected payoff with respect to the risky asset allocation, and then mitigates the moral hazard problem. Section 4.1 and Section 5.3 illustrate this result under discrete and continuous distributions of risk returns, respectively. When historical returns indicate high risk, the algorithm reduces its recommendations, thereby reducing managers' benefits. Consequently, the platform can force the manager's risk allocation to any desired level and further optimize the aggregate expected payoff for investors. Proverbially speaking, the platform effectively controls the market participants and interactions by leveraging "the algorithm's hand."

A key difference here, compared with recommendation algorithms on consumption platforms and social media, is the uncertainty of historical performance. Historical signals have varying degrees of informativeness, making it necessary to consider the reliability of the corresponding information structures in different situations. Consider an over-risky alternative portfolio with an overlapping return range to the target portfolio. In this case, the algorithm cannot be fully confident in distinguishing the manager's risk choice when they observe a realized return that falls within the overlap. In order to counter such uncertainty, the algorithm tends to be under-recommendation on such overlap. We also find in Section 4.1 that the algorithm compensates by over-recommendation when they observe informative signals from the non-overlapping return range of the target portfolio.

Due to this trade-off, some investors receive suboptimal recommendations, update their beliefs incorrectly and experience a loss of welfare. Under- and over-recommendations effectively

³Although active managers may be subject to disclosure rules (e.g., mutual fund filings), in practice these are often delayed and incomplete.

constitute an information rent paid by investors. Section 4.2 illustrates that investors with high risk aversion levels are affected first because of over-recommendation. In addition, the inevitable information rent incentivizes us to analyze the effectiveness of the optimal algorithms under different contracts, as detailed in Section 4.3. In particular, we discover that the expected investor payoff takes an inverted U-shaped form with respect to increasing management fees under optimal algorithm designs. This sheds light on the joint design of algorithms and contracts, a feature absent in recommendation algorithms across other scenarios.

The recommendation algorithm is more than an information gatekeeper. In Section 4.4, we compare the baseline model with several information structures. Even if investors are fully informed about their type and the portfolio’s historical performance, they fail to mitigate the moral hazard of managers. The reason is that they fairly never get themselves over-exposed to risk and never fall into ex-post inefficiency in any sub-games. Therefore, the population cannot generate punishment or compensation, resulting in a lack of commitment power. Section 4.5 compares recommendation algorithms and ranking systems. The focus of recommendation algorithms is on the fact that un-recommended funds are unobservable and that signals are private and personalized. This is fundamentally different from the model of ranking systems (e.g., [Huang et al., 2020](#)).

The rest of this paper is organized as follows. Section 2 introduces the general baseline model of this paper, including the timeline and the objectives of the players. Section 3 provides the fundamental necessary conditions for the equilibrium algorithm under the general setting. In order to simplify the analysis, we develop the detailed implications of the optimal algorithm with a discrete return in Section 4. Section 5 develops a framework for determining recommendation algorithms under continuous risk-return distributions. We establish the existence and uniqueness of the optimal algorithm in $W^{1,p}$, $1 < p \leq +\infty$, and characterize the optimal algorithm via variational inequalities. Section 6 concludes. Proofs, additional results and institutional background are collected in the Appendix.

Related literature. Our paper contributes to the growing literature on the impact of financial technologies on the asset management industry. Financial markets have become highly institutionalized ([Buffa et al., 2022](#)). Asset management platforms aggregate investments and increase adoption by introducing various technologies. These include convenient access to centralized information on fund rankings (e.g., [Huang et al., 2020](#); [Evans and Sun, 2021](#); [Ben-David et al., 2022](#); [Huang et al., 2022](#); [Hong et al., 2024](#)), and robo-advisors that offer personalized portfolio designs (e.g., [D’Acunto et al., 2019](#); [Loos et al., 2020](#); [Capponi et al., 2022](#)). Adopting the perspective that platforms have become intermediaries (e.g., [Stoughton et al., 2011](#); [Cookson et al., 2021](#)), we explore how they connect a large population of retail investors with delegated investment agencies.

In particular, we focus on the novel usage of personalized recommendation algorithms on

these platforms, which are endogenous and influence the behavior of platform participants. According to [Capponi et al. \(2022\)](#), while investors often misjudge their risk aversion, robo-advisors can identify and communicate accurate preferences through interactive adjustments. We extend this idea by linking the algorithm’s risk aversion identification ability with its effectiveness in coordinating the manager with the investors.

Meanwhile, empirical evidence shows that new technologies have an impact on participants’ behavior. The integration of daily consumption and investment activities has increased investors’ risk-taking behavior ([Hong et al., 2020](#)), flow of information amplifies the influence of attention-induced trading (e.g., [Kaniel and Parham, 2017](#); [Barber et al., 2022](#)), thereby incentivizing fund managers’ risk chasing for greater visibility ([Hong et al., 2024](#)). Our theoretical framework provides insights into how fintech platforms can leverage technology to mitigate this two-sided over-risk-taking phenomenon and guide proper trading behavior.

We also closely relate to the literature on asset management contracts by demonstrating that recommendation systems can effectively address the agency problem inherent in simple contracts. The inevitable agency problems of simple linear and limited-liability contracts have been frequently highlighted in the literature (e.g., [Innes, 1990](#); [Palomino and Prat, 2003](#)), particularly with regard to generating risk-taking incentives ([Stoughton, 1993](#); [Lee et al., 2019](#)).⁴ [Li and Tiwari \(2009\)](#) solves the problem of moral hazard in risk choices by using an option-type bonus fee with an appropriate benchmark. However, as emphasized by [D’Acunto and Rossi \(2021\)](#), a practical challenge is that there is a preference for offering simpler contracts to minimize the risk of operational errors by non-professional households. The realistic context also indicates that platforms generally have limited authority over the adjustment of contracts between investors and delegated managers. In this paper, we show that the automated and personalized recommendation algorithm can successfully address moral hazard of risk allocations under simple contracts, thereby equipping platforms with a powerful tool to strengthen their intermediary role.

This idea of using technology to influence contract enforcement shares a similar spirit with [Cong and He \(2019\)](#), which explores how blockchain technology can increase the range of variables that can be included in contracts, thus finding a niche in finance for the function of blockchain and smart contracts. From a broader financial theory perspective, optimal algorithm design explores a novel interplay between contract and information design — a promising combination studied in corporate finance (e.g., [Azarmsa and Cong, 2020](#); [Szydlowski, 2021](#); [Luo, 2021](#)).

The critical role of recommendation algorithms enriches the literature on using commitment mechanism to empower buyers in transactions. In a bilateral trade, [Roesler and Szentes \(2017\)](#) show that buyers can influence sellers’ pricing strategies by acquiring incomplete information.

⁴Existing literature widely studies asset management contracts in many respects, e.g., [He and Xiong \(2013\)](#); [Parlour and Rajan \(2020\)](#) consider contracts that incentivize manager’s efforts; [Buffa et al. \(2022\)](#) consider avoiding unskilled managers and impacts on market efficiency. Here we focus on the context of guarding against risk-taking.

Ichihashi and Smolin (2023) allow the buyer’s information to depend on the price. They prove that recommendation algorithms can safeguard total consumer surplus against personalized price discrimination. Our study focuses on the moral hazard problem in purchasing financial services, which particularly features noisy signals and uncertain payoffs.⁵ Low-quality information becomes critical for balancing investor welfare and managerial incentives in mutual fund market. This complements the empirical evidence of (Li et al., 2017), which shows that in mutual fund investment, retail investors have less information and have lower capacity to analyze information compared to institutional investors. Therefore, they face more ambiguity. The algorithm provides individual investors with incomplete information via recommendations, thus alleviating frictions in the fund market.

In terms of information transmission, the recommendation algorithm contributes to the large literature on information gatekeepers (Baye and Morgan, 2001). Many studies have examined the various motives and functions of platforms that strategically modify search results (e.g., Armstrong and Zhou, 2011; Hagiu and Jullien, 2011; Inderst and Ottaviani, 2012; De Corniere and Taylor, 2019; Zhou, 2020; Teh and Wright, 2022). Recent contributions of Bergemann and Bonatti (2024) emphasize that the consumer-seller platform exploits consumer data to increase its bargaining power with sellers. We share a similar spirit with distinct features that the algorithm also utilizes noisy information from managers, and ultimately aims to eliminate moral hazard.

2 The Model

Consider a two-period economy where investors with heterogeneous risk aversion enter into a contract with a fund manager via a platform. The manager’s limited liability induces moral hazard, potentially overexposing investors to market risk. In this paper, we assume that the platform targets a large user base and therefore attempts to protect investor welfare. Despite being unable to modify the contracts between investors and the manager, the platform is involved in the matching process by designing fund recommendation algorithms.

2.1 Setup

Assets and Fund manager. There is a risky asset and a risk-free asset. Without loss of generality, the risk-free return R_f is normalized to zero. The risky return R_t , $t = 1, 2$ is independently and identically distributed across two periods, following $R_t \sim G$, where G has a strictly positive density g over its support $[\underline{R}, \bar{R}]$ with $\mathbb{E}[R_t], \text{Var}[R_t] < \infty$. A risk-neutral manager designs a portfolio by determining the share of allocation to risky assets, $x \in [0, 1]$. Then the portfolio return

⁵Other researchers have studied commitment in mutual fund investments from different perspectives. For example, Huang et al. (2020) focus on shaping the market reputation in repeated games.

$R_{pt} = (1 - x)R_f + xR_t = xR_t$, $t = 1, 2$. We introduce two key assumptions to capture practical frictions. Firstly, neither the investors nor the platform observe the x directly. Typically, hedge fund and alternative investment managers are subject to limited disclosure requirements, and even active mutual fund managers only disclose their holdings at cyclical intervals (e.g. quarterly) and with delays. Therefore, we assume that x is the manager's private information, while the platform and investors could only have information about the underlying risky asset, i.e., they only know the distribution G . Secondly, the manager's allocation should be predictable over time. In practice, the allocation is usually persistent due to factors such as managers' consistent investment habits, asset preferences and beliefs, and the costs associated with making sharp adjustments to a portfolio over a short period of time. For simplicity, we assume that x remains constant throughout the two periods. A more general setting would allow for some variation, such as $x_2 = x_1 + \zeta$, where historical allocation only serves as a noisy information about the future allocation. Under this setting, the algorithm and the investors account for such additional uncertainty, while our main results and implications remain unchanged.

The manager sells the fund on the platform at $t = 1$. The financial payoff comes from the limited-liability delegated asset management contract, $\phi(r) = \max\{\alpha r, 0\} + \beta$, where the first term is a performance fee proportion $\alpha \geq 0$, and $\beta \geq 0$ is a fixed management fee. The manager may also be incentivized by personal benefits, such as becoming an attention-grabbing star with high performance, which has an asymmetric effect on the manager's utility in relation to the fund's gains and losses. This setting is analogous to the private benefits received by entrepreneurs when they succeed in financing (e.g., [Szydlowski, 2021](#)). The manager's expected utility reads

$$\mathbb{E}[u_M(R_{p2})] = q \left[\underbrace{\mathbb{E}[\phi(R_{p2})]}_{\text{financial payoff}} + \gamma \underbrace{\mathbb{E}[\max\{R_{p2}, 0\}]}_{\text{personal benefit}} \right],$$

where q is the total sales, $\gamma \geq 0$ is the personal benefit factor. The asymmetric revenue structure gives rise to agency problems: if the manager disregards the impact on sales, they have an incentive to fully allocate funds to risky assets. This is an inevitable consequence of a limited liability contract. As emphasised in [Palomino and Prat \(2003\)](#), such a structure prevents investors from selling returns to managers in exchange for their expected value, thereby exacerbating the misalignment of incentives between managers and investors.

Investors. A unit continuum of investors have heterogeneous risk aversion. They are indexed by their type a , where an a -type investor has \$1 to invest and decides whether to invest in the fund

at $t = 1$ based on the expected quadratic utility over the terminal return:

$$\mathbb{E}[u_1(R_{p2})] = \mathbb{E}[R_{p2} - \phi(R_{p2})] - \frac{1}{2}a\mathbb{E}[(R_{p2} - \phi(R_{p2}))^2],$$

The distribution F of type a over the population has a strictly positive density f over its support $[\underline{a}, \bar{a}]$, $0 < \underline{a} < \bar{a}$. As a baseline assumption, we consider that investors are unaware of their type a and unable to search for an appropriate portfolio independently. Instead, they only decide whether to invest in the fund after receiving a recommendation. When no recommendation is received, investors cannot observe a specific fund in the market. This assumption reflects the lack of financial expertise among platform users, including biased behavioral perceptions, a lack of information, and self-unawareness. See, for example [Capponi et al. \(2022\)](#). This friction becomes particularly relevant when platforms expand financial inclusion and attract inexperienced investors. In Section 4.4, 4.5 and Appendix B, we explore alternative information structures where investors (i) know a exactly, and (ii) are aware of the fund even without receiving a recommendation. These alternative cases are important because they reinforce the idea that investors' lack of (or biased) knowledge about their own type creates an opportunity for the algorithm to establish commitment power.

Platform and algorithm. The platform leverages its ability to collect data on investors' risk aversion and the fund's historical performance. It can implement a recommendation algorithm that delivers *personalized recommendation signals*, and this algorithm is *publicly* known to investors and the manager. Specifically, an algorithm is a function $m : [\underline{a}, \bar{a}] \times [\bar{R}, \underline{R}] \rightarrow [0, 1]$. For any pair of a and r_{p1} , the algorithm recommends the fund to an a -type investor with probability $m(a, r_{p1})$.

Timeline. Formally, the timeline is as follows:

1. Fintech platform designs an algorithm m , publicly known.
2. Nature draws investors' type a .
3. Manager designs a fund which generates a historical return R_{p1} .
4. Platform privately observes each investor's risk aversion a . With probability $m(a, R_{p1})$, a -type investors observe the recommendation and R_{p1} , then decide whether to contract with the recommended manager.
5. If contracted, investors and the manager earn $R_{p2} - \phi(R_{p2})$ and $\phi(R_{p2}) + \gamma\mathbb{E}[\max\{R_{p2}, 0\}]$, respectively.

2.2 Platform's Optimization and Solution Concept

The aim of this paper is to study whether and how the recommendation algorithm mitigates moral hazard and enhances social welfare under simple contracts. The solution concept is subgame-perfect equilibrium, and all integrals in this paper should be understood in the Lebesgue sense.

The platform's design of the recommendation algorithm maximizes investors' total expected payoff. We focus on the platform's user-side motivation based on two key considerations. From a regulatory perspective, robo-advisors are classified as fiduciaries under the Investment Advisers Act of 1940, which requires them to act in their clients' best interests (Capponi et al., 2022). From an incentive perspective, Xu and Yang (2023) emphasizes that the platforms aiming to maximize future revenue tend to be "consumer-oriented" since their business success depends heavily on past user satisfaction. For example, Uber employs technology-driven tools to mitigate the moral hazard of driver detours, thereby improving the passenger experience (Liu et al., 2021); More relevant to our scope, Yieldstreet, a tech-based customized asset management platform, collects earnings from investors rather than fund managers, making it naturally accountable to investors. See Appendix C for more institutional background.

The platform designs an algorithm m that restricts the manager's optimal allocation x in the equilibrium in order to maximize the aggregate investor expected payoff. Note that the risk-free rate is normalized to zero, formally we can write the problem as follows:

$$\max_{\substack{m: [a, \bar{a}] \times [\underline{R}, \bar{R}] \rightarrow [0, 1], \\ x \in [0, 1]}} \int_{\underline{R}}^{\bar{R}} \int_{\underline{R}}^{\bar{R}} \int_{\underline{a}}^{\bar{a}} \left[(xr_2 - \phi(xr_2)) - \frac{1}{2}a(xr_2 - \phi(xr_2))^2 \right] m(a, xr_1) dF(a) dG(r_2) dG(r_1) \quad (1)$$

subject to the following constraints

$$x \in \arg \max_{x'} \left\{ \int_{\underline{R}}^{\bar{R}} \int_{\underline{R}}^{\bar{R}} \int_{\underline{a}}^{\bar{a}} [\phi(x'r_2) + \gamma \max\{x'r_2, 0\}] m(a, x'r_1) dF(a) dG(r_2) dG(r_1) \right\}, \quad (2)$$

$$\int_{\underline{R}}^{\bar{R}} \int_{\underline{R}}^{\bar{R}} \int_{\underline{a}}^{\bar{a}} [\phi(xr_2) + \gamma \max\{xr_2, 0\}] m(a, xr_1) dF(a) dG(r_2) dG(r_1) \geq 0, \quad (3)$$

$$\int_{\underline{R}}^{\bar{R}} (xr_2 - \phi(xr_2)) - \frac{1}{2} \frac{\int_{\underline{a}}^{\bar{a}} am(a, xr_1) dF(a)}{\int_{\underline{a}}^{\bar{a}} m(a, xr_1) dF(a)} (xr_2 - \phi(xr_2))^2 dG(r_2) \geq 0, \forall r_1 \in \text{supp}(R_1). \quad (4)$$

Eq. (2) is the incentive compatibility (IC) constraint for the manager, meaning that the equilibrium allocation x will maximize the manager's expected financial payoff when using the corresponding recommendation algorithm. Eq. (3) is the manager's individual rationality (IR) constraint. It is naturally satisfied. Eq. (4) is the investors' IR constraint. Investors who receive recommendations form posterior beliefs about their risk aversion based on the publicly known algorithm. Given the observed historical return r_{p1} , their expected utility of investing in the rec-

ommended fund is given by

$$\begin{aligned} & \int_{\underline{R}}^{\bar{R}} \int_{\underline{a}}^{\bar{a}} (xr_2 - \phi(xr_2)) - \frac{1}{2}a (xr_2 - \phi(xr_2))^2 dF(a|\text{recommended}, xr_1) dG(r_2) \\ &= \int_{\underline{R}}^{\bar{R}} (xr_2 - \phi(xr_2)) - \frac{1}{2}\mathbb{E}[a|\text{recommended}, xr_1] (xr_2 - \phi(xr_2))^2 dG(r_2), \end{aligned}$$

which obtains the R.H.S. of (4). An investor who satisfies the IR constraint would prefer to contract when they receive a recommendation. In other words, this constraint limits the total sales influenced by the algorithm.

Note that Eq. (4) is equivalent to requiring the integral interior of the optimization problem (1) to be non-negative, that is

$$\int_{\underline{R}}^{\bar{R}} \int_{\underline{a}}^{\bar{a}} \left[(xr_2 - \phi(xr_2)) - \frac{1}{2}a (xr_2 - \phi(xr_2))^2 \right] m(a, xr_1) dF(a) dG(r_2) \geq 0.$$

Under the IR condition (4), there is no distinction between algorithmic recommendations and investor investments in the expression.

For convenience, we define the following notations:

$$\begin{aligned} \mu_+ &:= \int_0^{\bar{R}} r_2 g(r_2) dr_2, & A &:= (\alpha + \gamma)\mu_+, \\ k_1(x) &:= \int_{\underline{R}}^{\bar{R}} (xr_2 - \phi(xr_2)) dG(r_2), & k_2(x) &:= \int_{\underline{R}}^{\bar{R}} (xr_2 - \phi(xr_2))^2 dG(r_2). \end{aligned}$$

The following reasonable assumptions are made when solving the model:

Assumption 1.

1. The investor's utility $u_I(r)$ is increasing and concave: $1 - ar > 0$ for all $a \in [\underline{a}, \bar{a}]$ and all $r \in [\underline{R}, \bar{R}]$.
2. The contract does not prevent trading: $\mathbb{E}[R_t - \phi(R_t)] = \mathbb{E}[R_t] - \alpha\mu_+ - \beta > 0$.

Assumption 1.2 ensures that the contract costs do not become so high that the maximum expected return on the portfolio is lower than that of a risk-free asset. Under Assumption 1, we have that $\forall a \in [\underline{a}, \bar{a}]$,

$$\left. \frac{d\mathbb{E}[u_I(xR_t)]}{dx} \right|_{x=0} > 0, \quad \text{and} \quad \frac{d^2\mathbb{E}[u_I(xR_t)]}{(dx)^2} < 0.$$

3 Formalism Properties of the Optimal Algorithm

In this section, we analyze the properties of the optimal algorithm in a formalistic manner, i.e., we proceed to discover the key intuitive insights the algorithm should embody, without rigorously establishing its existence. The optimal algorithm belongs to a family of functions featuring a simple threshold that imposes a penalty for abnormal returns.

3.1 Threshold Algorithm

Inspired by the seminal work of [Ichihashi and Smolin \(2023\)](#), we consider the threshold algorithms. An algorithm m is a *threshold algorithm* if there exists a *threshold function* $\hat{a} : [\underline{R}, \bar{R}] \rightarrow [\underline{a}, \bar{a}]$ such that $m(a, r_{p1}) = \mathbb{1}(a < \hat{a}(r_{p1}))$. In other words, a threshold algorithm recommends the fund with probability 1 (0) if the investor's risk aversion is below (above) the threshold determined by historical returns.

Lemma 1. (Threshold algorithm.) *For any feasible algorithm m , there exists a threshold algorithm $\hat{m}$, under which the manager's expected payoff remains the same, whereas investors yield a (weakly) greater aggregated expected payoff than in the cases of m .*

According to Lemma 1, if there exists an optimal algorithm m^* , then we can always find a corresponding threshold algorithm $\hat{m}^*$ that ensures the investor's IR condition, the manager's IC condition and the manager's IR condition are all satisfied, and the investor's expected utility does not decrease. This suggests that the investor-optimal algorithm can be found within the set of threshold algorithms. In particular, for any given risky portfolio, investors with lower risk aversion always have higher expected utilities. Therefore, any recommended investor should have lower risk aversion than any unrecommended investor; otherwise, the total welfare could be increased by exchanging their recommendation states. Consequently, Lemma 1 suggests that if an optimal algorithm exists, it essentially determines a recommendation quota and then issues recommendations sequentially according to investors' risk aversion.

We can then represent the platform's problem (1) in terms of the fraction q of recommended investors determined by the threshold $\hat{a}$. Because the probability density function of a is strictly positive, q increases strictly with $\hat{a}$ over $[\underline{a}, \bar{a}]$, ranging from 0 to 1. Let $q := \int_{\underline{a}}^{\hat{a}} 1 dF(a) = F(\hat{a})$, where $\hat{a}$ is determined by the realized historical return $r_{p1} = xr_1$. With $q(xr_1) : [\underline{R}, \bar{R}] \rightarrow [0, 1]$ and $\hat{a}(xr_1) = F^{-1}(q(xr_1))$, the equilibrium can be represented as (x, q) , and the platform's problem is rewritten as

$$\max_{\substack{q: [\underline{R}, \bar{R}] \rightarrow [0, 1], \\ x \in [0, 1]}} \int_{\underline{R}}^{\bar{R}} k_1(x) q(xr_1) - \frac{1}{2} \left(\int_{\underline{a}}^{F^{-1}(q(xr_1))} a dF(a) \right) k_2(x) dG(r_1) \quad (5)$$

subject to

$$x \in \arg \max_{x'} \left\{ (Ax' + \beta) \int_{\underline{R}}^{\bar{R}} q(x'r_1) dG(r_1) \right\} \text{ and} \quad (6)$$

$$k_1(x)q(xr_1) - \frac{1}{2} \int_{\underline{a}}^{F^{-1}(q(xr_1))} adF(a)k_2(x) \geq 0, \forall r_1 \in \text{supp}(R_1). \quad (7)$$

Intuitively, the algorithm penalizes aggressive investment by linking the total sales to the historical portfolio returns. This has the potential to address the incentive issue in contracts. When a fund manager over-allocates to risky assets, they obtain a greater expected payoff from contracted (or recommended) investors due to limited liability. However, this would also generate an abnormal historical return relative to proper risk exposure, causing algorithms to reduce the proportion of recommended investors.

We can further represent the IR constraint in a simpler form. Multiply the both sides of (7) by $q(xr_1)$ and focus on the left hand side. The derivative w.r.t. $q(xr_1)$ is $[k_1(x) - 1/2F^{-1}(q(xr_1))k_2(x)]$ and is strictly decreasing w.r.t. $q(xr_1)$. Also note that (7) is equal when $q(xr_1) = 0$. We can then define $\bar{q}(x)$ as

$$\bar{q}(x) := \sup \left\{ q \in [0, 1] \mid k_1(x)q - \frac{1}{2} \int_{\underline{a}}^{F^{-1}(q)} adF(a)k_2(x) \geq 0 \right\},$$

and the IR constraint is equivalent to

$$q(xr_1) \leq \bar{q}(x), \quad \forall r_1 \in [\underline{R}, \bar{R}]. \quad (8)$$

Briefly, the applicable recommendation fraction q has an upper envelope, as investors would not buy if they find the platform over-delivers signals.

3.2 Non-Monotonic Algorithm

The primary principal-agent problem here is the manager's tendency to invest excessively in risky assets, which is driven by their expected utility that increases monotonically with the x . In practice, high ranking based on high returns generates a huge incentive for fund managers, pushing them to be more risk-chasing (Hong et al., 2024). The algorithm is designed to change this monotonicity by influencing $q(\cdot)$. To do so, $q(\cdot)$ should somehow sacrifice its monotonicity and relate to the risk-return distribution. The following proposition describes this observation.

Lemma 2. (Failure of monotonic algorithms.) *If the fraction $q(\cdot)$ of recommendation characterized by an algorithm is weakly increasing in $r \in [\underline{R}, \bar{R}]$, then the algorithm induces $x^* = 1$ in equilibrium.*

Intuitively, if $q(\cdot)$ weakly increases with r , then the algorithm provides the same incentives as the contract $\phi(\cdot)$ to the manager. Consequently, the manager will fully invest in risky assets in order to maximize the expected return. In this case, the algorithm fails to bind the manager's over-exposure to risk, although it could still prevent investors with negative expected utility under $x = 1$ from entering the market. An important implication is that the platform should reduce recommendations when historical returns are unusually high. Since $q(\cdot)$ can uniquely characterize a threshold algorithm, we will also refer to $q(\cdot)$ as the algorithm in the following sections.

More generally, the algorithm punishes abnormal returns that should be impossible under an equilibrium allocation. Therefore, we can further characterize a feasible form of the optimal algorithm as stated in the proposition below.

Proposition 1. (Equivalent cutoff algorithms.) *For any equilibrium (x^*, q^*) , there is an equilibrium $(x^*, \hat{q})$ which generates the same expected payoffs for the investors and manager as (x^*, q^*) does. Specifically, $\hat{q}$ takes a form of a "cutoff algorithm" where*

$$\hat{q}(r; x) := \begin{cases} q(r), & r \in \text{supp}(xR_t); \\ 0, & \text{otherwise.} \end{cases}$$

On the one hand, $\hat{q}$ implements a greater penalty than q^* once the realized return exceeds $\text{supp}(x^*R_t)$. On the other hand, $\hat{q}$ imposes no additional penalty in equilibrium x^* . This ensures that the equilibrium investor utility at x^* remains unchanged and that the $(x^*, \hat{q})$ pair still satisfies the IC constraint. In other words, the difference between q and $\hat{q}$ does not affect the reach and any quantitative nature of the equilibrium.

In what follows, we consider the existence of the equilibrium (x, q) and analyze the implications of the optimal algorithm in the form of $\hat{q}$ without loss of generality.

4 Optimal Algorithm under Discrete Return

This section develops the key ideas of this paper in the context of discrete return. Section 5 considers the case with continuous risk return. Here, we assume that a follows a uniform distribution, and the $\text{supp}(R_t) = \{\underline{R}, 0, \bar{R}\}$, reflecting the states "down", "flat", and "up", with probabilities $p(\underline{R})$, $p(0)$, and $p(\bar{R})$, respectively.

Given the discrete distribution, the optimization problem (5) can be transformed into solving for the allocation x^* and the three points $\hat{q}(x^*R_t)$ according to Lemma 1. Therefore, the existence of an investor-optimal algorithm is guaranteed through convex optimization on a compact set.

4.1 Moral Hazard Mitigation

We divide the optimization problem into two stages: (i) Given any target risk allocation x , find the optimal algorithm that will achieve the highest investor welfare among all feasible algorithms that reach the target x in equilibrium. (ii) Compare the resulting optimal welfare across different x , then determine the optimal risk allocation that maximizes total investor welfare. By Lemma 1, Proposition 1 and the equivalent investor's IR condition (8), the optimization problem given x is as follows:

$$\begin{aligned} & \max_{\hat{q}(x\underline{R}), \hat{q}(x\overline{R}), \hat{q}(0)} k_1(x) [\hat{q}(x\underline{R})p(\underline{R}) + \hat{q}(0)p(0) + \hat{q}(x\overline{R})p(\overline{R})] \\ & - \frac{1}{2}k_2(x) \left[\left(\int_{\underline{a}}^{F^{-1}(\hat{q}(x\underline{R}))} a dF(a) \right) \hat{q}(x\underline{R}) + \left(\int_{\underline{a}}^{F^{-1}(\hat{q}(0))} a dF(a) \right) \hat{q}(0) + \left(\int_{\underline{a}}^{F^{-1}(\hat{q}(x\overline{R}))} a dF(a) \right) \hat{q}(x\overline{R}) \right]. \end{aligned} \quad (9)$$

subject to

$$(Ax + \beta) [\hat{q}(x\underline{R})p(\underline{R}) + \hat{q}(0)p(0) + \hat{q}(x\overline{R})p(\overline{R})] \geq (A + \beta)\hat{q}(0)p(0). \quad (10)$$

$$0 \leq \hat{q}(xr_1) \leq \bar{q}(x), \forall r_1 \in \{\underline{R}, 0, \overline{R}\}. \quad (11)$$

Under the cutoff algorithm (defined in Proposition 1), if the manager deviates from the target value of x , they can only be recommended when the historic return equals zero. Therefore, once a deviation occurs, the manager will only deviate to $x' = 1$, and the IC constraint (6) is equivalent to (10). Since the constraint conditions are always satisfied by $(q(x\underline{R}), q(x\overline{R}), q(0)) = (0, 0, 0)$, feasible solutions always exist for any target x . The problem can be solved using the lagrangian multipliers. Given x , if a solution satisfies $\hat{q}(xr_1) \in (0, \bar{q}(x))$, then it can be written in the form of

$$\begin{aligned} \hat{q}(x\underline{R}) &= F \left(\frac{k_1(x) + \lambda(Ax + \beta)}{k_2(x)/2} \right), \\ \hat{q}(x\overline{R}) &= F \left(\frac{k_1(x) + \lambda(Ax + \beta)}{k_2(x)/2} \right), \\ \hat{q}(0) &= F \left(\frac{k_1(x) + \lambda A(x - 1)}{k_2(x)/2} \right), \end{aligned}$$

where $\lambda \geq 0$ is the multiplier of IC constraint (10).

To show the trade-off between scenarios of an algorithm, one could consider the ideal social planner's ex-post recommendation, which would involve recommending all investors with non-negative expected utilities. The social planner's recommendation can also be understood as the

solution to the optimization problem (9) without IC condition (10). The solution is

$$q_{FB}(x) = F(k_1(x)/(k_2(x)/2)). \quad (12)$$

Intuitively, moral hazard arises from the manager's advantage in hiding the allocation x , whereas the historical performance serves as information to infer x . In this three-point case, the platform correctly obtains x once the realized return $r_1 \neq 0$. Therefore, with probability $1 - p(0)$, the platform clearly knows x and delivers recommendations to the population accordingly. However, a "flat" realized price would blur all possible allocations. The key to mitigating moral hazard is therefore to introduce a penalty in the uninformative flat case, so that the platform is expected to narrow its recommendation delivery conservatively in order to protect investor welfare. But how can enough expected sales be generated to achieve ex-ante incentive compatibility for the manager? As compensation, the platform slightly expands its recommendations when it has information on x , even including investors with insufficient risk tolerance. This results in slight welfare losses, essentially an information rent paid to the manager.

Figure 1 visualizes the optimal algorithm $\hat{q}^*$ and optimal allocation x^* in a simulation. Panel (a) compares the optimal algorithm $\hat{q}^*$ and q_{FB} given x^* . When the historical return is zero, the algorithm recommends a lower probability than the first best to prevent the manager from deviating at a point where the algorithm cannot infer the x . Conversely, when the historical return is positive, the algorithm recommends a higher probability than the first-best, thereby imposing a more effective constraint on the manager's behavior. The similar phenomenon can also be seen in Panel (c) for different values of the x .

4.2 Information Rent

In this subsection we discuss more detailed properties of the information rents paid to the manager in our model. Firstly, consider the impact of the risk allocation x on the information rents. Denote the lower bound of perfect implementation as $\underline{x} := p(0) - (1 - p(0))\beta/A$. We will explain this constant later in this subsection. We also make the following assumption:

Assumption 2. *The manager is not willing to fully allocate in risk-free assets: $(A + \beta)p(0) > \beta$.*

Proposition 2. *(Algorithms and information rents.)*

(i) *(Without information rents.)* When the target equilibrium allocation $x^* \geq \underline{x}$, the optimal algorithm reaches the first best, mitigating moral hazard without paying information rent, $\hat{q}^*(x^* \underline{R}) = \hat{q}^*(0) = \hat{q}^*(x^* \bar{R}) = q_{FB}(x^*)$;

(ii) *(With information rents.)* When the target equilibrium allocation $x^* < \underline{x}$, the optimal algorithm

MITIGATING MORAL HAZARD THROUGH ALGORITHMS

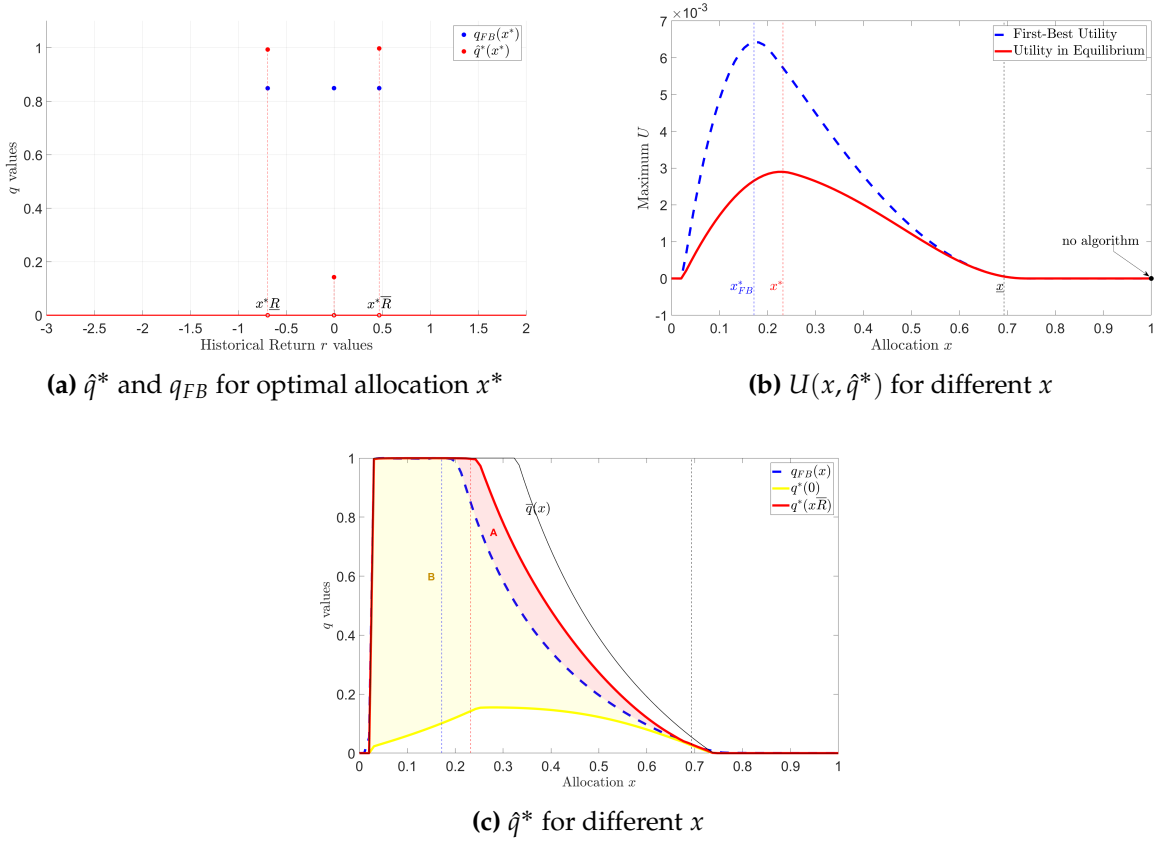


Figure 1: Optimal Algorithm $\hat{q}^*(\cdot)$ and Allocation x^*

Notes: Figure 1 illustrates an example of the optimal algorithm $\hat{q}^*$ and optimal allocation x^* . The parameters are chosen as follows: $\alpha = 0.01$, $\beta = 0.003$, $\gamma = 0.2$, $a \sim U[0.15, 0.5]$, and the support of R_t is $\{-3, 0, 2\}$ with corresponding probabilities 0.1, 0.7, and 0.2, respectively. Panel (a) compares the optimal algorithm $\hat{q}^*$ and q_{FB} given x^* . Compared with the first-best, the algorithm suffers from under-recommendation when $R_{p1} = 0$ and over-recommendation when $R_{p1} \in \{x^* \underline{R}, x^* \bar{R}\}$. Outside $\text{supp}(x^* R_1)$, the algorithm sets the recommendation probability to 0 as stated in Proposition 1. Panel (b) shows the expected utility of investors under optimal incentive-compatible and first-best algorithms, given different values of x . Panel (c) shows the optimal incentive-compatible and first-best algorithms for different values of x . Area A and B represent the over-recommended and under-recommended investor populations, respectively, resulting in the welfare gap in Panel (b).

strategically make recommendations to pay information rents and ensure IC condition,

$$\min\{\hat{q}^*(x^* \underline{R}), \hat{q}^*(x^* \bar{R})\} \geq q_{FB}(x^*) \geq \hat{q}^*(0).$$

When $q_{FB}(x^*) \in (0, 1)$, the inequalities hold strictly.

Proposition 2 highlights that $\underline{x}$ is a crucial threshold of the exposure to risk. When an algorithm targets an allocation x that exceeds this threshold, it mitigates moral hazard without paying any

information rent, acting as a de facto social planner. Economically, $x \geq \underline{x}$ is equivalently to

$$Ax + \beta \geq (A + \beta) p(0),$$

where the left hand side is the manager's expected payoff for good behavior in one deal, i.e., allocating x risky assets that align with the platform's target. The right hand side is the opportunity cost of being good, i.e., the maximum expected payoff of from deviating towards excessive risk-taking. In this case, the manager would choose a full allocation of risky assets and have a $p(0)$ probability recommendation. Thus $x \geq \underline{x}$ guarantees the incentive compatibility without additional rent. Otherwise, the platform requires additional effort to enforce the incentive constraint by penalizing the flat case and subsidizing other cases in terms of recommendation counts. In this case, the algorithm is designed to control the behavior of the fund manager based on all realized values of the return.

Panel (b) of Figure 1 illustrates the expected aggregate investor utility, which is represented by an inverted U-shaped curve. In the case of no algorithm, the only equilibrium is $x = 1$ and investors receive zero expected utility. In other words, algorithmic intervention removes the manager's full control over allocation, optimizing investor's utility in spite of potential information rent. Additionally, Panel (b) also shows the under-performance of the algorithm relative to the ideal case when $x < \underline{x}$, due to the required information rent and corresponding ex-post errors.

In this economy, a proportion of investors pay all the information rents. As Figure 1 Panel (c) shows, the ideal ex-ante recommendation q_{FB} is naturally determined by x and is independent of historical performance. Yet the algorithm's recommendation varies depending on the period-1 state. When the fund exhibits a flat state, the algorithm's recommendation scale is insufficient compared to q_{FB} . That is, some investors who were objectively eligible to invest are not recommended, resulting in foregone welfare gains, as depicted in Area B. Conversely, when the risky asset generates a non-flat state, the algorithm recommends an additional population with negative expected payoffs, as depicted in Area A.⁶

Intuitively, a lower x will require a more costly $\hat{q}^*$, so the welfare loss from deviating from the q_{FB} forms a trade-off with respect to the magnitude of the x . Consequently, one can expect x^* to exceed the optimal risk exposure at the same recommendation probability in equilibrium.

Proposition 3. (Comparison to the first-best situation with no moral hazard.)

1. For any targeted allocation x , the ex-ante expected recommendation fraction (with moral hazard) is

⁶Note that these investors voluntarily follow the recommendation guaranteed by the IR constraint. They achieve non-negative expected payoffs based on their subjective posterior risk aversion.

weakly lower than the first-best recommendation fraction q_{FB} :

$$\sum_{r \in \text{supp}\{R_t\}} p(r) \hat{q}^*(xr) \leq q_{FB}(x)$$

2. For an optimal algorithm $\hat{q}^*(\cdot)$ that realizes the targeted allocation x (with moral hazard), the first-best solution (with no moral hazard) prefers a lower x' under the same recommendation structure:

$$x \geq \sup \left\{ \arg \max_{x'} \left\{ k_1(x') \sum_{r \in \text{supp}\{R_t\}} p(r) \hat{q}^*(xr) - \frac{1}{2} \sum_{r \in \text{supp}\{R_t\}} \left(\int_a^{F^{-1}(\hat{q}^*(xr))} adF(a) \right) k_2(x') \right\} \right\}.$$

In particular, when $x \in (0, \underline{x})$, the inequalities hold strictly.

Proposition 3 shows the deviations in expected participation and risky asset allocation from the first-best solution. The reason underlying this is that algorithmic designs must account for information rent costs, as outlined in Proposition 2. In order to offset the monotonicity of the manager's utility with respect to risk exposure through total sales, the expected number of recommended investors must be reduced. Furthermore, to ensure the manager is incentive compatible, the algorithm concedes in risk exposure x^* .

The two parts of Proposition 3 can be unified as follows: The algorithm uses a lower investor participation rate to ensure that the targeted x satisfies the IC conditions. Under the threshold algorithm, the expected risk aversion levels of these participating investors are lower, meaning their optimal investment should actually be higher.

4.3 Algorithms under Different Contracts

Proposition 2 also implies the interplay between the algorithm and contract. The underlying logic arises from the manager's payoff structure: it hinges on the likelihood of being recommended and the expected returns once recommendation. The algorithm determines the former, while the contract determines the latter. Then the variation in contract design affects the manager's compromise on the algorithm when making allocation decisions.

In precise, the performance fee rate α and fixed management fee β jointly affect the threshold $\underline{x}$. A greater α decreases $\underline{x}$, resulting in a narrower range for the algorithm to achieve zero information rent, because it increases the incentive of higher returns, making the penalty from reduced recommendation less important. This is intuitive, as the performance fee with limited liability constitutes the origin of the principal-agent problem.

Consider the fixed management fee, β . Proposition 2 illustrates that a larger β allows the algorithm to achieve zero information rent at a broader range of equilibrium allocation x . Because

the management fee is independent of the portfolio performance, but solely depends on successful contract. This therefore becomes an incentive to align with the objectives of the platform's algorithmic design, and to avoid penalty in recommendation scales.

Figure 2 presents a comparative static analysis on β . As Panel (a) and (b) show, when the management fee is low, the equilibrium allocation under optimal algorithm, x^* , is higher than x_{FB}^* , and the expected recommendation scale is lower. As β increases, x_{FB}^* remains relatively stable, while x^* decreases to align with x_{FB}^* , and the under-recommendation at the flat-price scenario is resolved. It implies that the moral hazard gradually diminishes, as the manager is more like to align with the algorithm and earn the remarkable management fee. In particular, the social planner's solution can be achieved by the optimal algorithm when β is sufficiently high,⁷ i.e., $x^* > \underline{x}$ as outlined in Proposition 2.

To further analyze the investor welfare affected by the management fee, we decompose the expected aggregate ex-ante investor utility,

$$\mathbb{E}[u_I(x, \hat{q}^*(x))] = \underbrace{k_1(x)q_{FB}(x) - \frac{1}{2}k_2(x) \int_{\underline{a}}^{F^{-1}(q_{FB}(x))} a dF(a)}_{\text{First-best utility, } U(x, q_{FB})} - \underbrace{\left[k_1(x)\Delta_q(x) - \frac{1}{2}k_2(x)\Delta_a(x) \right]}_{\text{Utility Loss from Moral Hazard, } U(x, q_{FB}) - U(x, \hat{q}^*)},$$

where first term is the first-best payoff under x , and the second term represents the utility loss (e.g., over- and under-recommendation) to make x incentive-compatible. The loss is reflected as the deviations relative to the social planner's solution, including the expected recommend probability, $\Delta_q(x)$, and the expected collective risk aversion $\Delta_a(x)$ of recommended investors,

$$\begin{aligned} \Delta_q(x) &= \underbrace{-p(0)(q_{FB}(x) - \hat{q}^*(0))}_{\text{Under-recommendation}} + \underbrace{\sum_{r \in \{\underline{R}, \bar{R}\}} p(r)(\hat{q}^*(xr) - q_{FB}(x))}_{\text{Over-recommendation}}, \\ \Delta_a(x) &= \underbrace{-p(0) \int_{F^{-1}(\hat{q}^*(0))}^{F^{-1}(q_{FB}(x))} a dF(a)}_{\text{Under-recommendation}} + \underbrace{\sum_{r \in \{\underline{R}, \bar{R}\}} p(r) \int_{F^{-1}(q_{FB}(x))}^{F^{-1}(\hat{q}^*(xr))} a dF(a)}_{\text{Over-recommendation}}. \end{aligned}$$

The impact of the management fee β on investor welfare is twofold: a higher β enables the algorithm to better influence the manager's decisions, reducing the information rent to pay. On the other hand, it directly reduces investor wealth as a fixed charge. With the above decomposition, the former impact is reflected solely in the deviation from the first-best solution, i.e., $U(x^*, q_{FB}) - U(x^*, \hat{q}^*)$, while the latter also enters $U(x^*, q_{FB})$.

As Figure 2 (c) shows, when β is relatively low, the advantage of increasing β is evident (al-

⁷The threshold of a sufficiently-high β is (about) 0.0035 under the parameter choice of Figure 2.

MITIGATING MORAL HAZARD THROUGH ALGORITHMS

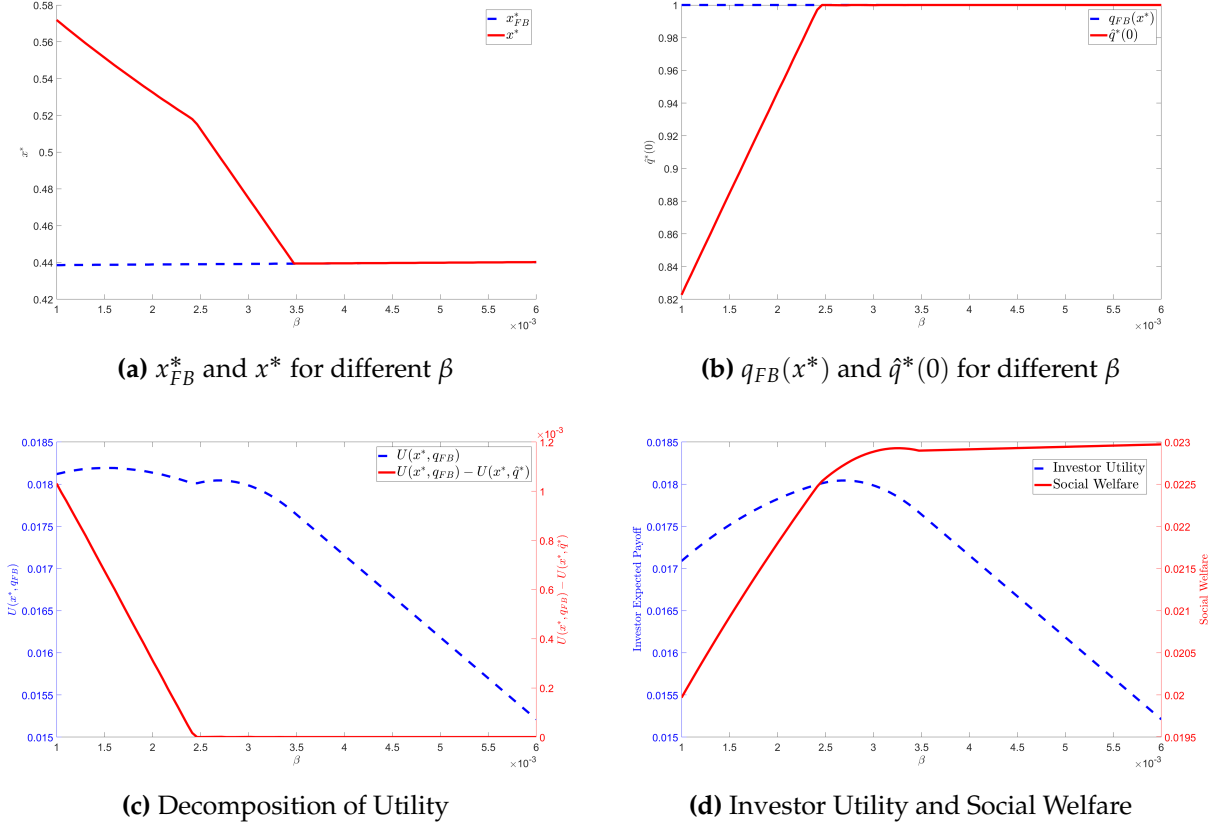


Figure 2: Comparative static analysis: management fee β in contract

Notes: Figure 2 illustrates the a comparative static analysis of the algorithm $\hat{q}^*$, allocation x^* and investor utility with respect to the management fee β . The parameters are chosen as follows: $\alpha = 0.02$, $\gamma = 0$, $a \sim U[0.5, 1]$, and the support of R_t is $\{-1, 0, 1\}$ with corresponding probabilities 0.1, 0.7, and 0.2, respectively. By comparing with with the First Best, it can be seen that as β rises, moral hazard diminishes and the algorithm $\hat{q}^*$ and the allocation x^* move closer to the First Best (see Panels (a) and (b)). Since an increase in β also directly results in a loss of investor utility, the utility initially increases with β but then declines (see Panels (c) and (d)).

though not fully offset the direct charge as $U(x^*, q_{FB})$ appears a decreasing trend), and the welfare gradually converges to the first-best case. When the optimal algorithm reaches the targeted equilibrium without information rent, the higher β only imposes costs, making $U(x^*, q_{FB})$ decreases linearly in β .

Combined these two forces, the investor utility exhibits an inverted U-shaped curve, as shown in Panel (d). This non-monotonicity suggests a space for *a jointly optimal design of the algorithm and contract*. Additionally, since β represents a mere transfer payment from investors to fund managers, the reduction in wealth effect is not accounted in the total social welfare calculation. As shown in Panel (d), the social welfare (the aggregate expected utility of investors and managers) increases due to the mitigation of the principal-agent problem.

4.4 Algorithm Serves as A Commitment Mechanism

A more fundamental question is: why the algorithm can successfully manipulate the equilibrium allocation and thereby mitigate the moral hazard? This subsection uncovers its critical role as a commitment mechanism. Specifically, we consider different timeline and information structures and show that even if the investors are as informed as the algorithm (know their own types as well as the portfolio historical performance), they fail to mitigate the fund manager's moral hazard.

Consider several alternative timelines as plotted in Figure 3: (i) Blind investment, where investors do not adopt a platform, but meet the fund manager by chance. Therefore, the investors have no information. (ii) Investment on fund distribution platforms, where the platform offers information about the funds without providing recommendations. That is, the investors know the historical return R_1 but do not know their risk aversion a . (iii) Investment experts. The investors know both their types and all the historical performance information.

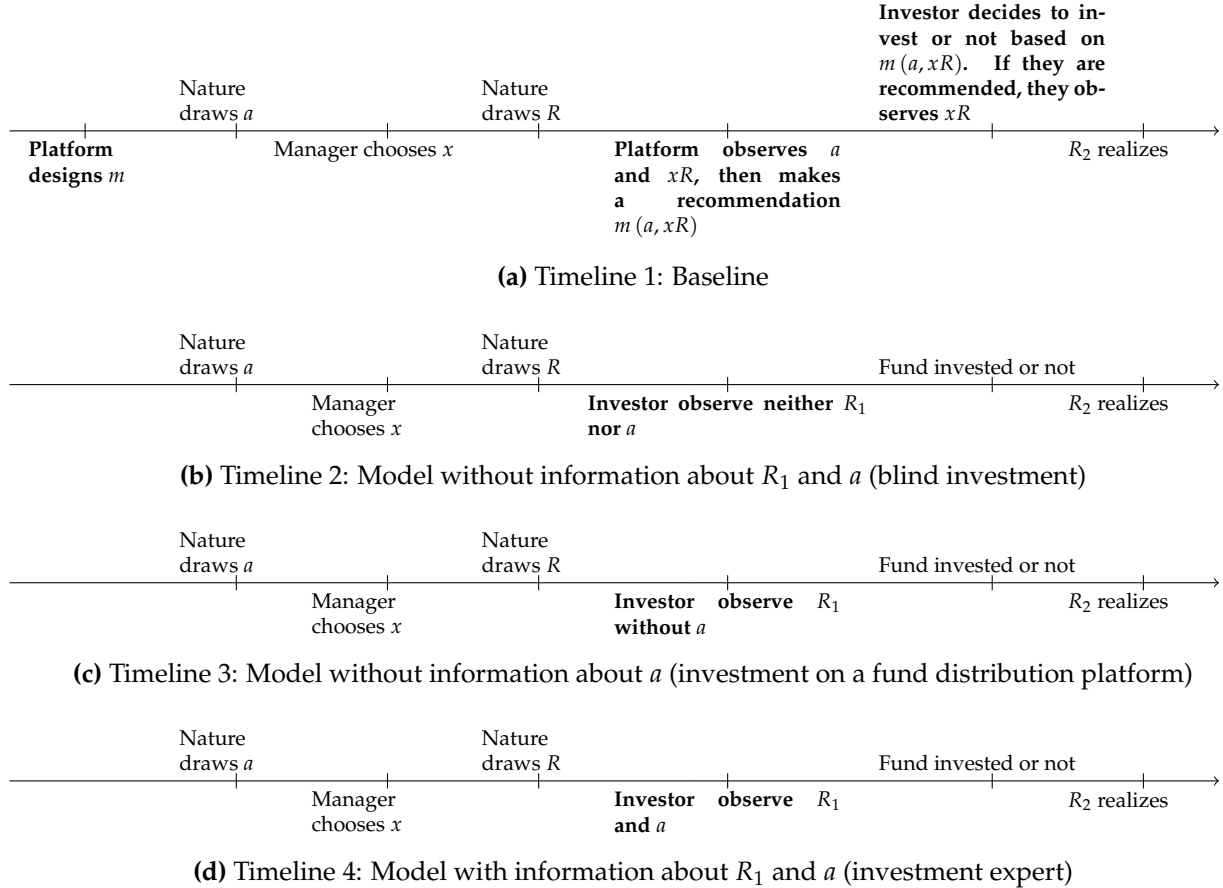


Figure 3: Alternative Timeline and Information Settings

We compare the total investor welfare under these alternative cases to the baseline. Appendix B provides formal analysis on each case, including proposition derivations and intuitions. We

fix the distributions of risk aversion and the risk return, then visualize the expected payoffs for investors at different risk aversion levels under each setting, as shown in Figure 4. A blind investment, as shown in the red dashed line, always results in a full risk taking ($x^* = 1$), and the manager takes away all the investor welfare. The green solid line shows the case of a fund distribution platform. The equilibrium reaches a slightly lower risk allocation ($x^* = 0.918$), as the investors could observe the historical return and partially infer the manager’s decision. However, they may be still over-exposed to risk due to limited awareness of their own risk preferences — particularly among highly risk-averse individuals, a group that lies at the heart of financial inclusion concerns. In Appendix B, we further show that the results are similar to this case even if the platform provides additional information, such as historical Sharpe ratios and more detailed information.

Investment experts, shown by the purple line of Figure 4, appear well-informed enough to protect themselves: when investors know both their types and the historical return, just as the algorithm does, they are never exposed to negative expected payoffs. However, the population fails to achieve a fairly low x .⁸ Only those with low risk aversion (shown in area A) choose to invest, excluding a large share of the population from participating in the delegated investment. The underlying intuition is: when every investor refuses to invest in situations with ex-post inefficiency, no one ends up paying the information rent. Since each investor makes decisions individually, the population as a whole is unable to penalize the manager’s risk-chasing behavior through a coordinated reduction in total sales. This lack of coordination, in a sense, exhibits a “curse of shrewdness” and fails to create a commitment power. In contrast, the blue line shows the baseline with an algorithm: it achieves an investor-optimal equilibrium $x^* = 0.515$. The piecewise pattern around $a = 4$ reflects the presence of information rent. Notably, the resulting aggregate expected payoff — roughly corresponding to area B — is greater than any other settings, and financial inclusion is substantially expanded.

The above discussion highlights the function of algorithm: it coordinates investor behavior to generate commitment power, thus maximizing aggregate investor welfare. Revisit its unique role relative to (interacted with) contracts. Given simple contracts that consider only future performance without historical records, the algorithm affects the business by collecting information and deciding signal delivery, effectively enabling functionalities of a series of complex contracts (including both historical and future conditions), and further determining the valid contract parties.

⁸Under the parameter choices of Figure 4, the resulting equilibrium allocation $x = 1$, while it is possible to reach an equilibrium with $x < 1$.

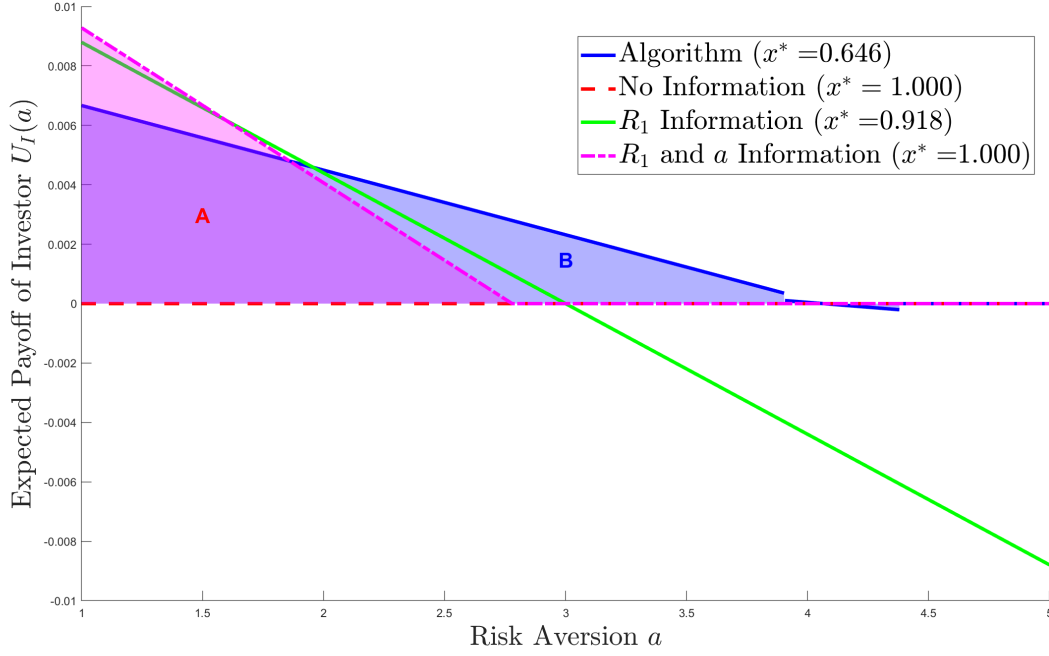


Figure 4: Expected payoff under different risk aversion a and information structures.

Notes: Figure 4 illustrates the distribution of investors' expected utility in equilibrium under four different information structures (ranked by risk aversion a). The parameters are chosen as follows: $\alpha = 0.1$, $\beta = 0.0015$, $\gamma = 0.5$, $a \sim U[1, 5]$, and the support of R_t is $\{-0.2, 0, 0.2\}$ with corresponding probabilities 0.1, 0.7, and 0.2, respectively. The blue solid line corresponds to the baseline model, the red dashed line represents the case where the investor has no information (blind investment), the green solid line corresponds to the case where the investor can observe R_1 (investment on a fund distribution platform), and the pink dashed line represents the case where the investor can accurately observe both R_1 and a (investment experts). It can be seen that allowing investors to access coarse information about their own a improves their overall welfare, as Area B is larger than Area A.

4.5 Algorithm and Ranking Systems

In addition to the role as a commitment mechanism, this subsection uncovers the other resulting critical role played by the algorithm, i.e., a private information gatekeeper, which is significantly distinct from ranking systems. Specifically, ranking systems (e.g., Morningstar ratings) aim to help investors compare funds and identify suitable investment targets. In contrast, recommendation algorithms start from investor heterogeneity: given a fund, they determine which investors are suitable for it. This motivation provides a tractable bridge between mechanism design and financial inclusion.⁹ A key difference is that the ratings are publicly known,¹⁰ while the algorithm delivers private signals according to investors' characteristics. In practice, the two are not mutually

⁹They both process historical performance data and somehow serve similar classification functions. For instance, an extremely high allocation x may lead the algorithm to implicitly classify the fund as "high-risk" and thus reduce recommendation.

¹⁰With a rating system, investors get to know a list of funds, at least the top funds. This suggests the crucial influence of ratings: they generate investors' attention to top funds, leading to risk chasing to hit the ranking (Hong et al., 2024).

exclusive, e.g., the platform can even publish ratings and deliver personalized recommendation simultaneously.

We consider the case when there are both recommendation algorithms and fund ratings. With the existence of publicly known ratings, investors always know the fund and thus can invest even without receiving recommendation. The platform's problem is formulated as: $x \in [0, 1]$,

$$\max_{q: [\underline{R}, \bar{R}] \rightarrow [0, 1]} k_1(x)q(xr_1) - \frac{1}{2} \left(\int_{\underline{a}}^{F^{-1}(q(xr_1))} a dF(a) \right) k_2(x) dG(r_1)$$

subject to the IC and IR constraints

$$x \in \arg \max_{x'} \left\{ (Ax' + \beta) [q(x'R)p(\underline{R}) + q(0)p(0) + q(x'\bar{R})p(\bar{R})] \right\},$$

$$k_1(x) - \frac{1}{2} \frac{\int_{\underline{a}}^{F^{-1}(q(xr_1))} a dF(a)}{q(xr_1)} k_2(x) \geq 0, \forall r_1 \in \text{supp}(R_1), \quad (13)$$

$$k_1(x) - \frac{1}{2} \frac{\int_{F^{-1}(q(xr_1))}^{\bar{a}} a dF(a)}{1 - q(xr_1)} k_2(x) \leq 0, \forall r_1 \in \text{supp}(R_1). \quad (14)$$

The additional IR constraint (14) implies that investors only follows the algorithm's recommendation to reject the investment action when the posterior expectation is high enough. In particular, as investors are able to know the fund and its performance via public information, the IR constraint (14) rules out cases where they still invest in the fund given no recommendation received. Otherwise, the algorithm cannot remain its commitment power.

Similar to processing the baseline IR constraint, we multiply the both sides of (14) with $(1 - q(xr_1))$, and consider the left side. Its derivative w.r.t. $q(xr_1)$ is

$$-k_1(x) + \frac{1}{2} F^{-1}(q(xr_1)) k_2(x),$$

and is strictly increasing with $q(xr_1)$. Also note the equal sign holds when $q(xr_1) = 1$. Then we can define $\underline{q}(x)$ as

$$\underline{q}(x) := \inf \left\{ q \in [0, 1] \mid k_1(x)(1 - q) - \frac{1}{2} \int_{F^{-1}(q)}^{\bar{a}} a dF(a) k_2(x) \leq 0 \right\}.$$

Then the IR constraint (14) is equivalent to $q(xr_1) \geq \underline{q}(x)$ for any $r_1 \in [\underline{R}, \bar{R}]$. Proposition 4 indicates how the two IR constraints bind and interact with the equilibrium allocation x .

Proposition 4. *(Applicable range of the recommendation affected by the ranking system.) Suppose the contract parameters and the distribution of risk aversion satisfy $k_1(x) - 1/2 \underline{a} k_2(x) \geq 0$. Then there are*

cases where the upper (lower) envelope of the applicable recommendation, $\underline{q}(\bar{q})$, takes different values,

1. If $k_1(x) - \frac{1}{2} \int_{\underline{a}}^{\bar{a}} a dF(a) k_2(x) < 0$, $\underline{q}(x) = 0$ and $\bar{q}(x) \in (0, 1)$.
2. If $k_1(x) - \frac{1}{2} \int_{\underline{a}}^{\bar{a}} a dF(a) k_2(x) > 0$, $\underline{q}(x) \in (0, 1)$ and $\bar{q}(x) = 1$.

Proposition 4 highlights the importance of the population average (or the average belief about the) risk aversion levels. (i) when x is too large for the population average risk aversion, investors would not invest when receiving no recommendations, while the baseline IR constraint binds: if the signals were over-delivered, investors would not buy; (ii) importantly, the new binding scenario is that when x is below the population average risk tolerance, the lower limit binds, i.e., investors may be inclined to invest even without a recommendation.

Figure 5 visualizes this case in comparison with the baseline, exactly corresponds to the second scenario in Proposition 4. First, the algorithm still significantly forces the equilibrium allocation away from $x = 1$, and the investor-optimal x^* roughly equals to 0.3. However, this is not from a convex optimization, but bound by the new IR constraint: when the algorithm aims at a low risk exposure $x < 0.3$, it needs to pay a remarkable information rent to the manager. Then it has to be too conservative such that $q^*(0) < \underline{q}(x)$. Then the investors ignore the fact of not being recommended, and still invest. As a result, the algorithm fails to reach an equilibrium at x .

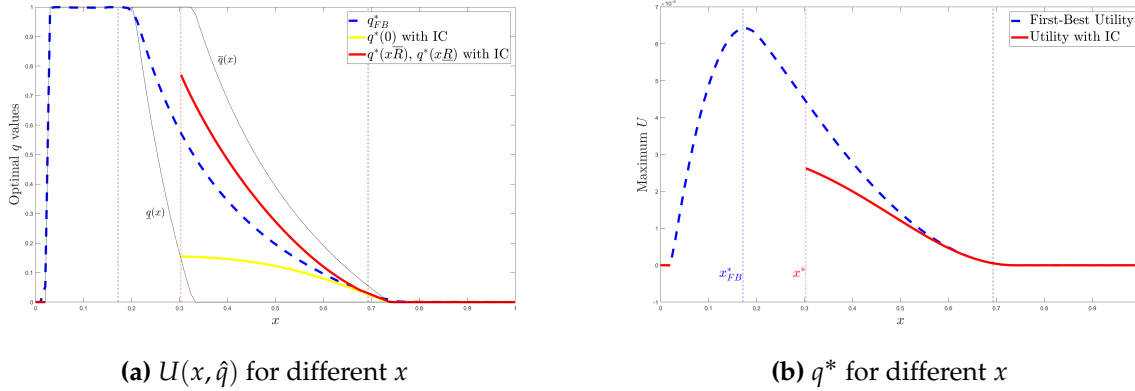


Figure 5: Investor's Payoff and $\hat{q}^*(\cdot)$ with constraint $q(xr_1) \geq \underline{q}(x)$

Notes: Figure 5 illustrates the optimal algorithm and utility for different x , when investors can observe all funds (i.e., the rating indicating whether to recommend purchasing). Compared to Figure 1, both the investor's utility and the algorithm are blank in the region below 0.3. This is because no incentive-compatible algorithm exists in this range—investors would invest even without a recommendation, causing fund managers to deviate. In this case, the optimal algorithm locks x^* at a higher level, leading to a lower expected utility for investors compared to Figure 1. The parameter choices are the same as Figure 1.

Compared to the baseline in Figure 1, the expected aggregate investor payoff decreases. This yields a counterintuitive implication: how can additional public information reduce social wel-

fare? Because once investors have access to alternative public signals, the algorithm’s recommendation becomes less influential in shaping their decisions. As a result, the platform’s ability to coordinate investor behavior — its commitment power — diminishes. This weaker coordination shifts the equilibrium risk allocation in favor of the manager. This finding aligns with [Hong et al. \(2024\)](#)’s empirical findings, where increased exposure in rating media is associated with increased exposure to risk. Given the platform can decide how to implement rating systems and recommendation algorithms, their interaction presents an important direction for future research in market information design.

5 Analysis under Continuous Distribution

In the previous section, uncertainty arises from the three possible future states. This simplifies the algorithm’s knowledge of the manager’s allocation into two cases: fully certain and fully unknown. In a more realistic continuous setting, each portfolio has a probability of yielding a continuum range of historical returns, albeit different to other portfolios. Therefore, any historical return fails to precisely infer the allocation. The algorithm is then expected to have strictly positive recommendation probabilities over a continuous interval of historical returns, rather than at discrete points as in Section 4.

In this section, we characterize the implications when $\text{supp}(R_t) = [\underline{R}, \bar{R}]$. The pre-determined algorithm infers the manager’s choices based on realized historical returns with varying confidence, and enforces different recommendations accordingly. One can then imagine that the previous implications still hold: the algorithm implements punishment, i.e., delivers conservative a recommendation, when receiving less informative and/or dangerous signals, and potentially pays an information rent. Ultimately, the algorithm optimally guides the manager to be incentive-compatible on lower risk exposures, mitigating moral hazard.

5.1 Existence of Optimal Algorithm

In a continuous setting, the existence of an optimal algorithm is not trivial because of the lack of the typical monotonicity. Proposition 2 implies that Helly’s selection theorem, commonly used in the literature on mechanism or contract design, cannot be applied directly. To define the existence problem rigorously, we constrain $q(\cdot)$ to a specific function space, denoted by $\mathcal{Q}$. Then the optimization problem (5) can be generalized as

$$\sup_{(x,q) \in D \cap [0,1] \times \mathcal{Q}} O(x,q),$$

$$D := \{(x, q) : q \in L^\infty((\underline{R}, \bar{R})), 0 \leq q(\cdot) \leq \bar{q}(x) \text{ and } (x, q) \text{ satisfies (6)}\},$$

$$\bar{q}(x) := \sup \left\{ q \in [0, 1] \mid k_1(x)q - \frac{1}{2} \int_{\underline{a}}^{F^{-1}(q)} a dF(a) k_2(x) \geq 0 \right\},$$

where $\bar{q}(x)$ represents the maximum possible fraction q under allocation x . That is, $q(xr_1) \leq \bar{q}(x)$, $\forall r_1 \in [\underline{R}, \bar{R}]$. It can be proved that $\bar{q}(x)$ is continuous with respect to x .

A reasonable choice of the function space $\mathcal{Q}$ ensures a certain sequential compactness of the feasible set, thus avoiding difficulties arising from not being able to constrain $q(\cdot)$ to be monotonic. For example, by constraining $q(\cdot)$ to a Lipschitz function space with the same constant L , the continuity of the objective function is mathematically guaranteed, as is the compactness of the feasible set. More generally, the specific choice of $\mathcal{Q}$ and the existence of model solutions are shown by the following theorem.

Theorem 1. (Existence.) Assume that F is supported on $[\underline{a}, \bar{a}]$ with continuous density function $f > 0$, where $-\infty < \underline{a} < \bar{a} < +\infty$. Assume also G has continuous density function $g > 0$ on $[\underline{R}, \bar{R}]$. Suppose $\bar{q}(\cdot)$ is continuous. Let $L > 0$ be a constant. If one of the following two conditions applies,

1. $\mathcal{Q} = \left\{ q \in C_b((\underline{R}, \bar{R})) : 0 \leq q(\cdot) \leq 1 \text{ and } \sup_{x \neq y} \frac{|q(x) - q(y)|}{|x - y|} \leq L \right\};$
2. $\mathcal{Q} = \left\{ q \in W^{1,p}((\underline{R}, \bar{R})) : 1 < p < \infty, 0 \leq q(\cdot) \leq 1 \text{ and } \|Dq\|_p \leq L \right\};$

Then there exists $(x^*, q^*) \in D \cap [0, 1] \times \mathcal{Q}$ such that $O(x^*, q^*) = \sup_{(x, q) \in D \cap [0, 1] \times \mathcal{Q}} O(x, q)$.

Remark 1. The notation and basic concept of the proof are outlined below. Those not engaged in detailed mathematical analysis may omit this Remark. Let $\Omega \subset \mathbb{R}$ be an open set. In the Theorem, $\|\cdot\|_{p;\Omega}$ denotes the $L^p(\Omega)$ norm with respect to Lebesgue measure. We omit the subscript Ω when the domain is clear (or not important) from the context.

By Du , we mean the distributional derivative of a L^1_{loc} function u . The spaces $W^{1,p}$, $1 \leq p \leq \infty$ are the Sobolev spaces of L^p functions with L^p first order distributional derivative. It can be shown that the space of Lipschitz continuous functions on bounded interval is just $W^{1,\infty}(\Omega)$. On the other hand, for function u of one real variable, $u \in W^{1,p}$, $1 < p < \infty$ implies u has a version $u' = u$ a.e. such that $u' \in C^{0,1-\frac{1}{p}}$. The Hölder class $C^{0,\alpha}$ consists of continuous functions such that $\sup_{x \neq y} \{|u(x) - u(y)| / |x - y|^\alpha\} < \infty$. Obviously, the space of Lipschitz functions is $C^{0,1}$. In this sense, the condition 1. and the condition 2. are similar and together they handle the $W^{1,p}$, $1 < p \leq \infty$ cases. For fundamental properties of the Sobolev spaces, we refer to [Adams and Fournier \(2003\)](#).

To prove Theorem 1, we adopt the direct method in the calculus of variations. To be more specific, we are going to show that under suitable topology the feasible set $D \cap [0, 1] \times \mathcal{Q}$ is se-

quentially compact and the objective function is at least upper semi-continuous. For the Lipschitz case, the desired compactness is guaranteed by Arzelà-Ascoli theorem. When the index $p < \infty$, $W^{1,p}$ spaces are reflexive and thus we rely on Banach-Alaoglu-Bourbaki theorem and Rellich-Kondrachov compactness theorem.

The intuition behind the constraint L in the space $\mathcal{Q}$ is that the derivative of $Q(\cdot)$ is not allowed to change drastically. This implies a large algorithmic design cost. This is similar to the cost of a certain energy functional, like $\int_{\underline{R}}^{\bar{R}} |q'(xr_1)|^2 dG(r_1)$ in some physics problems.

Solving for x^* and $q^*(\cdot)$ simultaneously presents certain challenges. To illustrate the solution, we divide the optimization problem into two stages. Firstly, we fix $x \in [0, 1]$ and find the solution $q^*(x)$ for the following optimization problem (15) and thus identifying the characteristics of the optimal algorithm.

$$\sup_{q \in \mathcal{Q} \cap D_x} O(q; x) \quad (15)$$

where $D_x := \left\{ q \in L^\infty((\underline{R}, \bar{R})) \mid 0 \leq q \leq \bar{q}(x) \text{ and } (x, q) \text{ satisfies (6)} \right\}$.¹¹ Secondly, we pin down the solution $(x^*, q^*(x^*))$ and the investor's utility, which allows us to comprehend the comprehensive impact of the algorithm on the principal-agent problem.

Here we observe the (partial) convexity of the objective function $O(\cdot, \cdot)$. Since it is useful hereafter, we call it a lemma:

Lemma 3. (Concavity of the objective function.) Suppose F is supported on $[\underline{a}, \bar{a}]$ with continuous density function $f > 0$, where $-\infty < \underline{a} < \bar{a} < +\infty$. Then, for any $x \in (0, 1]$, the objective function $O(x, q)$ is strictly concave with respect to q .

By Lemma 3, $q^*(x)$ is well-defined as demonstrated in the following proposition.

Proposition 5. (Uniqueness.) Given $x \in [0, 1]$, there exists a unique function $q^*(x)$ optimizing (15).

5.2 Solving for the Optimal Recommendation

We attempt to solve the optimal algorithm. The incentive constraint (6) can be rather complex for further analytical derivation. Here we alternatively propose a *local incentive constraint* (the first-order condition of (6) w.r.t. x) for potential solutions, and verify its satisfaction of the original

¹¹Since D_x includes $q(\cdot) \equiv 0$ for all x , D_x is not empty.

condition. The alternative constraint reads:

$$\underbrace{(\alpha + \gamma) \int_0^{\bar{R}} r_2 dG(r_2) \int_{\underline{R}}^{\bar{R}} q(xr_1) dG(r_1)}_{\text{Marginal expected payoff}} + \underbrace{\left[\beta + (\alpha + \gamma)x \int_0^{\bar{R}} r_2 dG(r_2) \right] \int_{\underline{R}}^{\bar{R}} q'(xr_1) r_1 dG(r_1)}_{\text{Marginal algorithmic penalty}} = 0, \quad (16)$$

where $\int_{\underline{R}}^{\bar{R}} q'(xr_1) r_1 dG(r_1)$ is well-defined according to Lebesgue's dominated convergence theorem. Denote (16) in a general form with functional I ,

$$\int_{\underline{R}}^{\bar{R}} I(r_1, q, q') dr_1 = 0.$$

We make two reasonable assumptions for sake of the solving process.

Assumption 3. Given x , for any $q \in \mathcal{Q}$,

$$\frac{\partial I}{\partial y}(r_1, q, q') - \frac{d}{dr_1} \left(\frac{\partial I}{\partial z}(r_1, q, q') \right) = A - (Ax + \beta)(r_1 g'(r_1) + g(r_1))$$

is not equal to zero a.e. in $[\underline{R}, \bar{R}]$.

Assumption 4. Given x , let $q^* \in D_x \cap \mathcal{Q}$ be the unique solution of the objective function. Fix any element $q_1 \in \mathcal{Q}$, $\exists q_2 \in \mathcal{Q}$ with

$$\frac{\int_{\underline{R}}^{\bar{R}} \frac{\partial I}{\partial y}(r_1, q^*, q^{*'}) (q_1 - q^*) + \frac{\partial I}{\partial z}(r_1, q^*, q^{*'}) (q'_1 - q^{*'}) dr_1}{\int_{\underline{R}}^{\bar{R}} \frac{\partial I}{\partial y}(r_1, q^*, q^{*'}) (q_2 - q^*) + \frac{\partial I}{\partial z}(r_1, q^*, q^{*'}) (q'_2 - q^{*'}) dr_1} < 0.$$

Assumption 3 is simply satisfied when $(r_1 g'(r_1) + g(r_1))$ is not constant.¹² Assumption 4 effectively assumes there exists an interior solution q^* , i.e., given x , $\exists r_1$ such that $q^*(xr_1) \in (0, \bar{q}(x))$, whilst the corner cases $\{0, \bar{q}(x)\}$ can be easily analyzed separately. With these two assumptions, we obtain a variational inequality as a necessary condition for the solution q^* of the optimization problem (15) given x .

Theorem 2. (Variational inequality as a necessary condition.) Given $x \in (0, 1)$, let $q^* \in D_x \cap \mathcal{Q}$ be the unique solution of the objective function. Under Assumption 3 and 4, there exists a real number λ s.t.

¹²Appendix Example 1 discusses the counter case where $r_1 g'(r_1) + g(r_1) = C$, i.e., $g(r_1) = C + C_1/|r_1|$. In particular, when G follows a uniform distribution, the algorithm always automatically achieves the first-best equilibrium with zero information rent.

$\forall q_1 \in \mathcal{Q}$,

$$0 \geq \int_{\underline{R}}^{\bar{R}} (q_1(xr_1) - q^*(xr_1)) \left[k_1(x) - \frac{1}{2}k_2(x)F^{-1}(q^*(xr_1)) + \lambda \frac{\beta}{x} + \lambda(A + \frac{\beta}{x})r_1 \frac{g'(r_1)}{g(r_1)} \right] dG(r_1) \\ + (q_1(xr_1) - q^*(xr_1)) \lambda(A + \frac{\beta}{x})r_1 g(r_1) \Big|_{\underline{R}}^{\bar{R}}. \quad (17)$$

Now we show how this necessary condition restricts the potential q^* to a specific formula. Consider that the set $U := \{r_1 \in (\underline{R}, \bar{R}) \mid 0 < q^*(xr_1) < \bar{q}(x)\}$ is open, and $C := \{r_1 \in [\underline{R}, \bar{R}] \mid q^*(xr_1) = 0 \text{ or } q^*(xr_1) = \bar{q}(x)\}$ is (relatively) closed. Fix any text function $v \in C_c^\infty(U)$. Then if $|\delta|$ is sufficiently small, $0 \leq q_1 := q^* + \delta v \leq \bar{q}(x)$, $q_1 \in \mathcal{Q}$. Thus (17) implies¹³

$$\int_U \delta v(xr_1) \left[k_1(x) - \frac{1}{2}k_2(x)F^{-1}(q^*) \right] dG(r_1) - \lambda \int_U A \delta v(xr_1) + (Ax + \beta)r_1 \delta v'(xr_1) dG(r_1) \leq 0.$$

This inequality is valid for both δ and $-\delta$. Therefore, the above inequality must have the equal sign. Because v has compact support in U , v is vanished near ∂U . By the integration by parts, we obtain

$$0 = \int_U v(xr_1) \left[k_1(x) - \frac{1}{2}k_2(x)F^{-1}(q^*(xr_1)) + \lambda \frac{\beta}{x} + \lambda(A + \frac{\beta}{x})r_1 \frac{g'(r_1)}{g(r_1)} \right] dG(r_1)$$

is valid for all $v \in C_c^\infty(U)$. Therefore,

$$0 = k_1(x)g(r_1) - \frac{1}{2}k_2(x)F^{-1}(q^*(xr_1))g(r_1) + \lambda \frac{\beta}{x}g(r_1) + \lambda(A + \frac{\beta}{x})r_1 g'(r_1) \quad \text{in } U. \quad (18)$$

Notably, when $U = (\underline{R}, \bar{R})$, we can also fix any text function $v \in C_c^\infty(\bar{U})$, where $\bar{U}$ is the closure of U . Then if $|\delta|$ is sufficiently small, $0 \leq q_1 := q^* + \delta v \leq \bar{q}(x)$ and so $q_1 \in \mathcal{Q}$ thus satisfies (17). Similarly, the inequality must have the equal sign. Together with (18), we obtain

$$0 = \lambda \delta (Ax + \beta) \left[(\bar{R}g(\bar{R})/x)v(x\bar{R}) - (\underline{R}g(\underline{R})/x)v(x\underline{R}) \right]. \quad (19)$$

Consider v such that $0 = v(x\underline{R}) < v(x\bar{R})$, there must be $\lambda = 0$.

So far, we draw the conclusion from Theorem 2 that $\exists \lambda \in \mathbb{R}$, s.t.¹⁴

$$q^*(xr_1) = F \left(\frac{k_1(x) - \lambda [\beta/x + (A + \beta/x)r_1 g'(r_1)/g(r_1)]}{k_2(x)/2} \right). \quad (20)$$

¹³Note that since $v \in C_c^\infty(U)$, $v(x\underline{R}) = v(x\bar{R}) = 0$.

¹⁴The expression $q^* = F(a^*)$ in (20) and (21) implicitly assumes $a^* \in (\underline{a}, \bar{a})$, while the other cases are relatively trivial, corresponding to the algorithm that never/always recommends to each investor.

In particular, if $U = (\underline{R}, \bar{R})$,

$$q^*(xr_1) = F\left(\frac{k_1(x)}{k_2(x)/2}\right). \quad (21)$$

5.3 Inverted U-shaped Algorithm

We pin down the return distribution for further analysis. Let the risk aversion of investors follow a uniform distinction, $a \sim U[0.15, 0.5]$, and the distribution of risk returns R_t be a truncated normal distribution supported on $[-3, 2]$, with $\mu = 0.5$ and $\sigma = 1.5$.¹⁵ In precise, the probability density function of R_t reads

$$g(r; \mu, \sigma, \underline{R}, \bar{R}) = \frac{1}{\sigma\sqrt{2\pi}} \frac{\exp\left(-\frac{1}{2}\left(\frac{r-\mu}{\sigma}\right)^2\right)}{\Phi\left(\frac{\bar{R}-\mu}{\sigma}\right) - \Phi\left(\frac{\underline{R}-\mu}{\sigma}\right)},$$

where $\Phi(\cdot)$ is the cumulative CDF of the standard normal distribution.

Figure 6 visualizes the optimal algorithm q^* solved from (20) that ensures the equilibrium $x = 0.4$.¹⁶ Panel (a) shows how the algorithm delivers recommendations over realized returns. First, the algorithm only delivers recommendation in a narrower support of observed returns, since highly abnormal returns are less likely to be the portfolio return with risk exposure $x \leq 0.4$. Second, the algorithm reduces recommendation when the historical return is higher. This aligns with the crucial intuition in Section 3.2: a high historical return may not be a good sign, as it may result from over-exposed to risk. Mechanically, the downward slope in Panel (a) reflects the increasing possibility of a too-large allocation, thereby aggregating more punishment. In addition, information rent is paid in this scenario. For example, when the recommendation amount goes beyond the first-best, $q_{FB}(x)$, there exists investors who are recommended and receive negative expected payoff.¹⁷

Recall the solving process. Theorem 2 only provides a necessary condition by alternatively satisfying the local incentive constraint—we need to verify that the IC constraint is satisfied. As Panel (b) shows, the algorithm successfully breaks the monotonicity of manager utility: the manager maximizes the expected payoff under the optimal algorithm, therefore is incentive compatible at $x = 0.4$, validating $q^*(x)$ to be an equilibrium algorithm. In particular, the manager's utility function becomes concave due to the non-monotonic algorithm. This means the algorithm effectively

¹⁵We use a truncated normal distribution for technically satisfying Assumption 1. Essentially, it could fully capture the intuitions of the risk return normally distributed over $(-\infty, +\infty)$: as we show, the algorithm would choose not to recommend if an extremely abnormal return was observed. It is somehow equivalent to presume the plausible returns to be distributed over a finite interval.

¹⁶Note that $x = 0.4$ may be not the optimal x^* that maximizes the aggregate expected investor utility. Essentially, any allocation can be achieved in equilibrium by designing a corresponding algorithm with potential information rent.

¹⁷The return interval with $q^* \leq q_{FB}(x)$ may also contain information rent, which is rather complex to decompose from the punishment, i.e., the recommendation amount may be lower if only accounting for the punishments of multiple possible allocations. While in the discrete case, the decomposition is clear since the allocation is correctly inferred.

transfers the investors' risk-aversion to the risk-neutral manager.

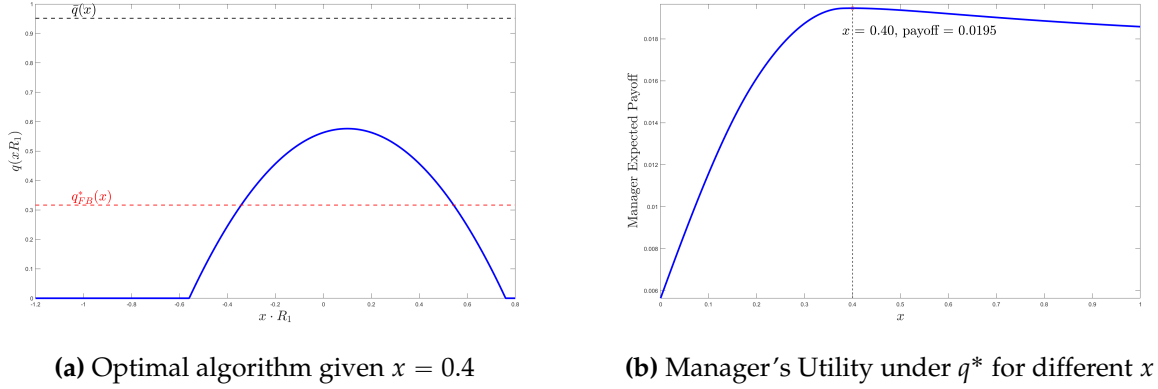


Figure 6: Optimal algorithm given $x = 0.4$ and Manager's expected payoff

Notes: Figure 6 illustrates how the algorithm incentivizes the fund manager to choose $x = 0.4$ rather than $x = 1$ in the continuous case. The parameters of the contract and the manager's personal benefit are set to $\alpha = \beta = 0.01$ and $\gamma = 0.2$. $a \sim U[0.15, 0.5]$. The risk return follows a truncated normal distribution supported on $[-3, 2]$, $\mu = 0.5$ and $\sigma = 1.5$. Panel (a) shows that the optimal algorithm is non-monotonic over $\text{supp}(xR_t)$ (with zero probability outside the support). Within the non-zero region, the algorithm is essentially quadratic, recommending with a higher probability than the first-best when returns are moderate, and a lower probability than the first-best when returns are extreme. Under the influence of this non-monotonic algorithm, panel (b) shows that the manager's utility function is no longer linear in x (based on the assumption of risk neutrality), but instead becomes a concave, non-monotonic function, reaching its maximum at $x = 0.4$.

Investor-optimal algorithm and information rent. Then we consider the equilibrium allocation and algorithm that maximize the aggregate expected investor payoff, where the interesting question is, similar to Section 4.2, does such optimum require an information rent? Note that whenever the algorithm does not induce $x^* = 1$, the effectiveness and operation of the algorithm rely on the distribution of returns, in terms of the multiplier λ and the elasticity of probability density functions, $(\partial/\partial(\ln r_1)) \ln(g(r_1))$.

First, when $\lambda = 0$, the elasticity does not matter, and the algorithm is simply a bang-bang form. The following Corollary draws implications, no information rent, and the necessary condition to achieve this scenario.

Corollary 1. *When the investor-optimal equilibrium yields a zero Lagrange multiplier, i.e., $\lambda^* = 0$, investors pay no information rent. In particular, The sufficient and necessary condition of $\lambda^* = 0$ is that, x^**

satisfies

$$\begin{aligned} x^* \in \arg \max_{x' \in [x^*, 1]} \left\{ (x' A + \beta) [G(x^* \bar{R}/x') - G(x^* \underline{R}/x')] \right\} \text{ and} \\ x^* \in \arg \max_{x'} \left\{ k_1(x') F \left(\frac{k_1(x')}{k_2(x')/2} \right) - \frac{1}{2} \left(\int_0^{k_1(x')/(k_2(x')/2)} \text{ad}F(a) \right) k_2(x') \right\}. \end{aligned} \quad (22)$$

Second, when $\lambda \neq 0$, the information rent has to be paid, and the elasticity of the density function directly determines the shape of the algorithm. Similar to contract and information design, algorithm design contributes to investor surplus improvement through the commitment power that may lead to ex-post inefficiencies.

Proposition 6. (Over- and under-recommendations in general cases.) *If $\lambda^* \neq 0$, there exists $r, r' \in [\underline{R}, \bar{R}]$ such that investors will be over- and under-recommended when historical returns are r and r' respectively. i.e.,*

$$\begin{aligned} \mathbb{E}[R_2 - \phi(R_2)|x^*] - \frac{1}{2} \hat{a}^*(r) \mathbb{E}[(R_2 - \phi(R_2))^2|x^*] < 0, \text{ and} \\ \mathbb{E}[R_2 - \phi(R_2)|x^*] - \frac{1}{2} \hat{a}^*(r') \mathbb{E}[(R_2 - \phi(R_2))^2|x^*] > 0, \end{aligned}$$

where $\hat{a}^*(r) = F^{-1}(q^*(r))$.

Similar to the discrete distribution case, Proposition 6 indicates that when $\lambda^* \neq 0$, there exist recommend investors who receive negative expected payoff and also unrecommended investors who could have positive expected payoff from investment. In addition, as discussed in Section 4.3, the contract and the algorithm interact with each other.

6 Conclusion

We develop a model of recommendation algorithm design in delegated investment. The intermediate platform serves investors who have limited knowledge about their risk aversion levels, aiming to mitigate fund managers' moral hazard in over risk-taking, particularly given the contractual environment unchanged. We show that predetermined automatic algorithms can effectively mitigate the principal-agent problem inherent in linear and limited-liability contracts. The core intuition is that the algorithm reshapes the information transmission and further affects the buyer party's entrance, which effectively generates commitment power. Specifically, the optimal algorithm is non-monotonic w.r.t. the fund's historical performance, thereby distorting the manager's utility function w.r.t. risk allocation. The manager has the incentive to hide behind noisy

signals. Therefore, the algorithm reduces recommendations when information is ambiguous and potentially compensates for informative signals. This generates an information rent paid by investors, facilitating trading and achieving Pareto improvement.

Although this paper focuses on risk incentives, the framework can be extended to designing algorithms for solving other principal–agent problems, such as managers’ efforts and information acquisition. See, for example [He and Xiong \(2013\)](#); [Huang et al. \(2020\)](#); [Buffa et al. \(2022\)](#). Our analysis may also inspire future research into many topics, such as optimal joint design with contracts, competition with multiple funds, and information design with both public fund ratings and personalized recommendations.

Furthermore, the powerful algorithm requires careful consideration of its regulation and purpose. The algorithm’s commitment power relies heavily on transparency, while as [Sun \(2024\)](#) points out, in reality, algorithms may not be fully transparent. Moreover, if the platform leans towards the manager side rather than the user base, the algorithm may establish a new principal–agent relationship that is detrimental to investor surplus. Overall, our paper reveals the extent to which algorithms can mitigate moral hazard in the context of digital finance.

References

- Adams, Robert A and John JF Fournier**, *Sobolev spaces*, 2 ed., Vol. 140 of *Pure and Applied Mathematics*, Elsevier, 2003.
- Armstrong, Mark and Jidong Zhou**, “Paying for prominence,” *The Economic Journal*, 2011, 121 (556), F368–F395.
- Azarmsa, Ehsan and Lin William Cong**, “Persuasion in relationship finance,” *Journal of Financial Economics*, 2020, 138 (3), 818–837.
- Barber, Brad M, Xing Huang, Terrance Odean, and Christopher Schwarz**, “Attention-induced trading and returns: Evidence from Robinhood users,” *The Journal of Finance*, 2022, 77 (6), 3141–3190.
- Baye, Michael R and John Morgan**, “Information gatekeepers on the internet and the competitiveness of homogeneous product markets,” *American Economic Review*, 2001, 91 (3), 454–474.
- Ben-David, Itzhak, Jiawei Li, Andrea Rossi, and Yang Song**, “What do mutual fund investors really care about?,” *The Review of Financial Studies*, 2022, 35 (4), 1723–1774.
- Bergemann, Dirk and Alessandro Bonatti**, “Data, competition, and digital platforms,” *American Economic Review*, 2024, 114 (8), 2553–2595.
- Berk, Jonathan B and Richard C Green**, “Mutual fund flows and performance in rational markets,” *Journal of political economy*, 2004, 112 (6), 1269–1295.

- Buffa, Andrea M, Dimitri Vayanos, and Paul Woolley**, “Asset management contracts and equilibrium prices,” *Journal of Political Economy*, 2022, 130 (12), 3146–3201.
- Capponi, Agostino, Sveinn Olafsson, and Thaleia Zariphopoulou**, “Personalized robo-advising: Enhancing investment through client interaction,” *Management Science*, 2022, 68 (4), 2485–2512.
- Cong, Lin William and Zhiguo He**, “Blockchain disruption and smart contracts,” *The Review of Financial Studies*, 2019, 32 (5), 1754–1797.
- Cookson, Gordon, Tim Jenkinson, Howard Jones, and Jose Vicente Martinez**, “Best buys and own brands: investment platforms’ recommendations of mutual funds,” *The Review of Financial Studies*, 2021, 34 (1), 227–263.
- Corniere, Alexandre De and Greg Taylor**, “A model of biased intermediation,” *The RAND Journal of Economics*, 2019, 50 (4), 854–882.
- D’Acunto, Francesco and Alberto G Rossi**, “Robo-advising,” in “The Palgrave Handbook of Technological Finance,” Springer, 2021.
- , **Nagpurnanand Prabhala, and Alberto G Rossi**, “The promises and pitfalls of robo-advising,” *The Review of Financial Studies*, 2019, 32 (5), 1983–2020.
- Evans, Lawrence C**, *Partial differential equations*, 2 ed., Vol. 19 of *Graduate Studies in Mathematics*, American Mathematical Society, 2010.
- Evans, Richard B and Yang Sun**, “Models or stars: The role of asset pricing models and heuristics in investor risk adjustment,” *The Review of Financial Studies*, 2021, 34 (1), 67–107.
- Hagiu, Andrei and Bruno Jullien**, “Why do intermediaries divert search?,” *The RAND Journal of Economics*, 2011, 42 (2), 337–362.
- He, Zhiguo and Wei Xiong**, “Delegated asset management, investment mandates, and capital immobility,” *Journal of Financial Economics*, 2013, 107 (2), 239–258.
- Hong, Claire Yurong, Xiaomeng Lu, and Jun Pan**, “FinTech adoption and household risk-taking: From Digital Payments to Platform Investments,” Technical Report, National Bureau of Economic Research 2020.
- , —, and —, “FinTech platforms and mutual fund distribution,” *Management Science*, 2024.
- Huang, Chong, Fei Li, and Xi Weng**, “Star ratings and the incentives of mutual funds,” *The Journal of Finance*, 2020, 75 (3), 1715–1765.
- Huang, Jennifer, Kelsey D Wei, and Hong Yan**, “Investor learning and mutual fund flows,” *Financial Management*, 2022, 51 (3), 739–765.
- Ichihashi, Shota and Alex Smolin**, “Buyer-Optimal Algorithmic Consumption,” Available at SSRN 4635866, 2023.
- Inderst, Roman and Marco Ottaviani**, “Competition through commissions and kickbacks,” *American Economic Review*, 2012, 102 (2), 780–809.
- Innes, Robert D**, “Limited liability and incentive contracting with ex-ante action choices,” *Journal*

- of Economic Theory*, 1990, 52 (1), 45–67.
- Kaniel, Ron and Robert Parham**, “WSJ Category Kings—The impact of media attention on consumer and mutual fund investment decisions,” *Journal of Financial Economics*, 2017, 123 (2), 337–356.
- Lee, Jung Hoon, Charles Trzcinka, and Shyam Venkatesan**, “Do portfolio manager contracts contract portfolio management?,” *The Journal of Finance*, 2019, 74 (5), 2543–2577.
- Li, C Wei and Ashish Tiwari**, “Incentive contracts in delegated portfolio management,” *The Review of Financial Studies*, 2009, 22 (11), 4681–4714.
- , —, and **Lin Tong**, “Investment decisions under ambiguity: Evidence from mutual fund investor behavior,” *Management Science*, 2017, 63 (8), 2509–2528.
- Liu, Meng, Erik Brynjolfsson, and Jason Dowlatabadi**, “Do digital platforms reduce moral hazard? The case of Uber and taxis,” *Management Science*, 2021, 67 (8), 4665–4685.
- Loos, Benjamin, Alessandro Previtero, Sebastian Scheurle, and Andreas Hackethal**, “Robo-advisers and investor behavior,” *Unpublished Working Paper*, 2020.
- Luo, Dan**, “Moral Hazard and the Corporate Information Environment,” *Available at SSRN* 3935008, 2021.
- Palomino, Frédéric and Andrea Prat**, “Risk taking and optimal contracts for money managers,” *RAND Journal of Economics*, 2003, pp. 113–137.
- Parlour, Christine A and Uday Rajan**, “Contracting on credit ratings: Adding value to public information,” *The Review of Financial Studies*, 2020, 33 (4), 1412–1444.
- Roesler, Anne-Katrin and Balázs Szentes**, “Buyer-optimal learning and monopoly pricing,” *American Economic Review*, 2017, 107 (7), 2072–2080.
- Stoughton, Neal M**, “Moral hazard and the portfolio management problem,” *The Journal of Finance*, 1993, 48 (5), 2009–2028.
- , **Yuchang Wu, and Josef Zechner**, “Intermediated investment management,” *The Journal of Finance*, 2011, 66 (3), 947–980.
- Sun, Jian**, “Algorithmic transparency,” *Unpublished Working Paper*, 2024.
- Szydlowski, Martin**, “Optimal financing and disclosure,” *Management Science*, 2021, 67 (1), 436–454.
- Teh, Tat-How and Julian Wright**, “Intermediation and steering: Competition in prices and commissions,” *American Economic Journal: Microeconomics*, 2022, 14 (2), 281–321.
- Xu, Wenji and Kai Hao Yang**, “Informational Intermediation, Market Feedback, and Welfare Losses,” *Unpublished Working Paper*, 2023.
- Zhou, Jidong**, “Improved information in search markets,” *Unpublished Working Paper*, 2020.

Appendices

A Derivation of Results

A.1 Proof of Lemma 1

The proof is following [Ichihashi and Smolin \(2023\)](#). Take any feasible algorithm m . For any fund with historical r_{p1} , let $q_m(r_{p1}) := \int_{\underline{a}}^{\bar{a}} m(a, r_{p1}) dF(a)$ denote the expected number of investors with recommendation under r_{p1} . Define a new algorithm $\hat{m}$ as $\hat{m}(a, r_{p1}) \equiv \mathbb{1}(a < F^{-1}(q_m(r_{p1})))$. At each r_{p1} , algorithm $\hat{m}$ recommends the fund with the same expected number of investors as m :

$$\int_{\underline{a}}^{\bar{a}} \mathbb{1}(a < F^{-1}(q_m(r_{p1}))) dF(a) = F(F^{-1}(q_m(r_{p1}))) = q_m(r_{p1}).$$

As a result, the fund manager earns the same profit under m and $\hat{m}$, both are

$$\mathbb{E}_{R_1}[q_m(xR_1)] \left[\mathbb{E}_{R_2}[\phi(xR_2)] + \gamma \mathbb{E}_{R_2}[\max\{xR_2, 0\}] \right].$$

Because the expected value of risk aversion a conditional on recommendation is lower under $\hat{m}$ than under m and $\partial \mathbb{E}_{R_2}[u_I(xR_2 - \phi(xR_2)); x] / \partial a < 0$, an investor who follows the recommendations of m would also follow those of $\hat{m}$, and the expected payoff for investors is higher under $\hat{m}$ than under r_{p1} .

A.2 Proof of Proposition 1

(i) According to the definition of $\hat{q}$, for any function $J(\cdot)$, given x^* , $\mathbb{E}_{R_1}[J(\hat{q}(x^*R_1))] = \mathbb{E}_{R_1}[J(q(x^*R_1))]$. Therefore, the expected payoffs of the investors and the manager are unchanged. Thus the investor's IR constraint still holds. (ii) Consider the IC condition. Since $\hat{q}$ induces potential additional penalties when $x' \neq x^*$ (when $x' < x$, the penalty would not trigger),

$$(Ax' + \beta) \int_{\underline{R}}^{\bar{R}} \hat{q}(x'r_1) dG(r_1) \leq (Ax' + \beta) \int_{\underline{R}}^{\bar{R}} q^*(x'r_1) dG(r_1).$$

Further by the incentive compatibility under equilibrium (x^*, q^*) and no extra penalty of $\hat{q}$ at x^* , we obtain

$$(Ax' + \beta) \int_{\underline{R}}^{\bar{R}} q^*(x'r_1) dG(r_1) \leq (Ax^* + \beta) \int_{\underline{R}}^{\bar{R}} q^*(x^*r_1) dG(r_1) = (Ax^* + \beta) \int_{\underline{R}}^{\bar{R}} \hat{q}(x^*r_1) dG(r_1).$$

Therefore, $(x^*, \hat{q})$ also satisfies the IC condition.

A.3 Proof of Proposition 2

For $x \in [0, 1]$, the algorithm design problem can be rewritten as

$$\max_{q: [\underline{R}, \bar{R}] \rightarrow [0, 1]} k_1(x)q(xr_1) - \frac{1}{2} \left(\int_{\underline{a}}^{F^{-1}(q(xr_1))} adF(a) \right) k_2(x) dG(r_1)$$

subject to the IC and IR constraint

$$x \in \arg \max_{x'} \left\{ (Ax' + \beta) [q(x'\underline{R})p(\underline{R}) + q(0)p(0) + q(x'\bar{R})p(\bar{R})] \right\}, \quad (\text{A1})$$

$$q(x) \leq \bar{q}(x). \quad (\text{A2})$$

All values of q on $[\underline{R}, \bar{R}]$ need to be determined. We can show that for any $(x, q(x))$ satisfying the IC constraint (A1), there exists $\hat{q}(x)$ defined by (A3),

$$\hat{q}(r; x) := \begin{cases} q(r), & \text{if } r \in \{x\underline{R}, 0, x\bar{R}\} \\ 0, & \text{otherwise,} \end{cases} \quad (\text{A3})$$

such that $(x, \hat{q}(x))$ satisfies the same IC constraint, and the investor's payoff under $(x, \hat{q}(x))$ is the same as that under $(x, q(x))$.

Therefore, the objective function can be further rewritten as

$$\begin{aligned} & \max_{q(x\underline{R}), q(x\bar{R}), q(0)} k_1(x) [\hat{q}(x\underline{R})p(\underline{R}) + \hat{q}(0)p(0) + \hat{q}(x\bar{R})p(\bar{R})] \\ & - \frac{1}{2} k_2(x) \left[\left(\int_{\underline{a}}^{F^{-1}(\hat{q}(x\underline{R}))} adF(a) \right) \hat{q}(x\underline{R}) + \left(\int_{\underline{a}}^{F^{-1}(\hat{q}(0))} adF(a) \right) \hat{q}(0) + \left(\int_{\underline{a}}^{F^{-1}(\hat{q}(x\bar{R}))} adF(a) \right) \hat{q}(x\bar{R}) \right]. \end{aligned}$$

Without the IC constraint, for any x , the optimal algorithm should recommend all investors whose payoff is non-negative under x . In this case, the first-best algorithm without the IC constraint is given by $q_{FB}(x) = k_1(x)/(k_2(x)/2)$ and we have

$$x_{FB}^* = \sup \left\{ \arg \max_{x'} \left\{ k_1(x') F \left(\frac{k_1(x')}{k_2(x')/2} \right) - \frac{1}{2} \left(\int_{\underline{a}}^{k_1(x')/(k_2(x')/2)} adF(a) \right) k_2(x') \right\} \right\}. \quad (\text{A4})$$

With the assumption $(A + \beta)p(0) > \beta$, under $\hat{q}$, the IC constraint (A1) is equivalent to

$$(Ax + \beta) [\hat{q}(x\underline{R})p(\underline{R}) + \hat{q}(0)p(0) + \hat{q}(x\bar{R})p(\bar{R})] \geq \max_{x'} (Ax' + \beta) \hat{q}(0)p(0) = (A + \beta) \hat{q}(0)p(0).$$

Note that if the manager deviates from x to x' , then they will not be recommended when the realization $r_1 \neq 0$. Consequently, given the algorithm, the probability of being recommended is

given by $\hat{q}(0)$ and the manager will only deviate towards $x' = 1$ to ensure that the expected return is maximized when they is recommended.

Given x , we have the following Lagrangian:

$$\begin{aligned}
 \mathcal{L}(\hat{q}(x\underline{R}), \hat{q}(0), \hat{q}(x\overline{R})) &= k_1(x) [\hat{q}(x\underline{R})p(\underline{R}) + \hat{q}(0)p(0) + \hat{q}(x\overline{R})p(\overline{R})] \\
 &- \frac{1}{2}k_2(x) \left[\left(\int_{\underline{a}}^{F^{-1}(\hat{q}(x\underline{R}))} adF(a) \right) p(\underline{R}) + \left(\int_{\underline{a}}^{F^{-1}(\hat{q}(0))} adF(a) \right) p(0) + \left(\int_{\underline{a}}^{F^{-1}(\hat{q}(x\overline{R}))} adF(a) \right) p(\overline{R}) \right] \\
 &+ \lambda \left[(Ax + \beta) [\hat{q}(x\underline{R})p(\underline{R}) + \hat{q}(0)p(0) + \hat{q}(x\overline{R})p(\overline{R})] - (A + \beta)\hat{q}(0)p(0) \right] \\
 &+ \eta_{\underline{R},1}\hat{q}(x\underline{R}) + \eta_{\underline{R},2}(\bar{q}(x) - \hat{q}(x\underline{R})) + \eta_{\overline{R},1}\hat{q}(x\overline{R}) + \eta_{\overline{R},2}(\bar{q}(x) - \hat{q}(x\overline{R})) + \eta_{0,1}\hat{q}(0) + \eta_{0,2}(\bar{q}(x) - \hat{q}(0)),
 \end{aligned} \tag{A5}$$

where $\lambda \geq 0$ and $\eta_{r,2} \geq 0$ are the multipliers from IC and IR constraints, and $\eta_{r,1} \geq 0$ is the multiplier from the definition of q .

According to the first-order condition of the Lagrangian, we have

$$\begin{aligned}
 F^{-1}(\hat{q}(x\underline{R})) &= \frac{k_1(x) + \lambda(Ax + \beta) + (\eta_{\underline{R},1} - \eta_{\underline{R},2})/p(\underline{R})}{k_2(x)/2}, \\
 F^{-1}(\hat{q}(x\overline{R})) &= \frac{k_1(x) + \lambda(Ax + \beta) + (\eta_{\overline{R},1} - \eta_{\overline{R},2})/p(\overline{R})}{k_2(x)/2}, \\
 F^{-1}(\hat{q}(0)) &= \frac{k_1(x) + \lambda A(x - 1) + (\eta_{0,1} - \eta_{0,2})/p(0)}{k_2(x)/2}.
 \end{aligned}$$

Plug $\hat{q}^*(x^*\underline{R}) = \hat{q}^*(0) = \hat{q}^*(x^*\overline{R}) = q_{FB}(x^*)$ into the IC constraint, we get

$$\begin{aligned}
 (Ax^* + \beta)[1 - p(0)] [q_{FB}(x^*) - \underline{a}] + A(x^* - 1)p(0) [q_{FB}(x^*) - \underline{a}] &\geq 0 \\
 \Rightarrow [(Ax^* + \beta)(1 - p(0)) + A(x^* - 1)p(0)] [q_{FB}(x^*) - \underline{a}] &\geq 0.
 \end{aligned}$$

Since $\max_r \{q^*(r)\} > 0$, the equilibrium x^* must ensure that the expected utility of the least risk-averse investor is non-negative, implying that $q_{FB}(x^*) \geq \underline{a}$. Then if

$$\begin{aligned}
 0 &\leq (Ax^* + \beta)(1 - p(0)) + A(x^* - 1)p(0) \\
 \Rightarrow x^* &\geq \frac{Ap(0) - (1 - p(0))\beta}{A} := \underline{x},
 \end{aligned} \tag{A6}$$

the IC constraint is satisfied. Additionally, since for a given x^* , $q_{FB}(x^*)$ maximizes the objective function and $q_{FB}(x^*) \leq q(x^*)$, then if $x^* \geq \underline{x}$, $\hat{q}^*(x^*\underline{R}) = \hat{q}^*(0) = \hat{q}^*(x^*\overline{R}) = q_{FB}(x^*)$ satisfies all constraints and maximizes the objective function, i.e., it is a solution. If $x^* < \underline{x}$, (A6) does not hold. Therefore, $\lambda > 0$, which further implies $\min\{\hat{q}^*(x^*\underline{R}), \hat{q}^*(x^*\overline{R})\} \geq q_{FB}(x^*) \geq \hat{q}^*(0)$.

A.4 Proof of Proposition 3

We primarily discuss the proof of the interior solution.

First, when $x^* \geq \underline{x}$ or $\lambda^* = 0$, the equation holds.

Secondly, when $x^* < \underline{x}$ and $\lambda^* > 0$, we have

$$\begin{aligned} \sum_{r \in \text{supp}\{R_t\}} p(r) \hat{q}^*(x^* r) &= \frac{k_1(x^*)}{k_2(2)/2} + (1 - p(0)) \frac{\lambda(Ax + \beta)}{k_2(x^*)/2} + p(0) \frac{\lambda A(x - 1)}{k_2(x)/2} \\ &= \frac{k_1(x^*)}{k_2(2)/2} + \frac{\lambda(Ax^* + \beta - p(0)(A + \beta))}{k_2(x^*)/2}. \end{aligned}$$

Given $\lambda > 0$ and $x^* < \underline{x}$, the second term is strictly negative, as a result we have

$$q^E := \sum_{r \in \text{supp}\{R_t\}} p(r) \hat{q}^*(x^* r) < \frac{k_1(x^*)}{k_2(2)/2}. \quad (\text{A7})$$

Given $\hat{q}^*(\cdot)$, the FOC of the Lagrangian (A5) with respect to x is

$$\begin{aligned} k'_1(x^*) q^E(x^*) - \frac{1}{2} k'_2(x^*) \left[p(0) \int_{\underline{a}}^{F^{-1}(\hat{q}^*(0))} adF(a) + \sum_{r \in \{\underline{R}, \bar{R}\}} p(r) \int_{F\underline{a}}^{F^{-1}(\hat{q}^*(xr))} adF(a) \right] + \lambda A q^E &= 0 \\ \Rightarrow k'_1(x^*) q^E(x^*) - \frac{1}{2} k'_2(x^*) \left[p(0) \int_{\underline{a}}^{F^{-1}(\hat{q}^*(0))} adF(a) + \sum_{r \in \{\underline{R}, \bar{R}\}} p(r) \int_{F\underline{a}}^{F^{-1}(\hat{q}^*(xr))} adF(a) \right] &< 0. \end{aligned}$$

The second line is because $\lambda > 0$ and $q^E > 0$. When

$$\tilde{x} := \sup \left\{ \arg \max_{x'} \left\{ k_1(x') \sum_{r \in \text{supp}\{R_t\}} p(r) \hat{q}^*(x^* r) - \frac{1}{2} \sum_{r \in \text{supp}\{R_t\}} \left(\int_{\underline{a}}^{F^{-1}(\hat{q}^*(x^* r))} adF(a) \right) k_2(x') \right\} \right\},$$

we have

$$k'_1(\tilde{x}) q^E(x^*) - \frac{1}{2} k'_2(\tilde{x}) \left[p(0) \int_{\underline{a}}^{F^{-1}(\hat{q}^*(0))} adF(a) + \sum_{r \in \{\underline{R}, \bar{R}\}} p(r) \int_{F\underline{a}}^{F^{-1}(\hat{q}^*(xr))} adF(a) \right] < 0.$$

A.5 Proof of Proposition 4

Recall that $\bar{q}(x)$ and $\underline{q}(x)$ are defined as below:

$$\bar{q}(x) = \sup \left\{ q \in [0, 1] \mid k_1(x) q - \frac{1}{2} \int_{\underline{a}}^{F^{-1}(q)} adF(a) k_2(x) \geq 0 \right\},$$

$$\underline{q}(x) = \inf \left\{ q \in [0, 1] \mid k_1(x)(1 - q) - \frac{1}{2} \int_{F^{-1}(q)}^{\bar{a}} adF(a)k_2(x) \leq 0 \right\}.$$

Given x , suppose $k_1(x) - \frac{1}{2}\underline{a}k_2(x) > 0$. We can observe that

$$k_1(x) \times 0 - \frac{1}{2} \int_{\underline{a}}^{F^{-1}(0)} adF(a)k_2(x) = 0, \quad (\text{A8})$$

$$k_1(x) \times (1 - 1) - \frac{1}{2} \int_{F^{-1}(1)}^{\bar{a}} adF(a)k_2(x) = 0, \quad (\text{A9})$$

$$\begin{aligned} \frac{dk_1(x)q - \frac{1}{2} \int_{\underline{a}}^{F^{-1}(q)} adF(a)k_2(x)}{dq} \Big|_{q=0} &= k_1(x) - \frac{1}{2} F^{-1}(q)k_2(x) \Big|_{q=0} \\ &= k_1(x) - \frac{1}{2}\underline{a}k_2(x) > 0, \end{aligned} \quad (\text{A10})$$

$$\begin{aligned} \frac{dk_1(x)(1 - q) - \frac{1}{2} \int_{F^{-1}(q)}^{\bar{a}} adF(a)k_2(x)}{dq} \Big|_{q=0} &= - \left[k_1(x) - \frac{1}{2} F^{-1}(q)k_2(x) \right] \Big|_{q=0} \\ &= - \left[k_1(x) - \frac{1}{2}\underline{a}k_2(x) \right] < 0, \end{aligned} \quad (\text{A11})$$

$$\frac{d^2 k_1(x)q - \frac{1}{2} \int_{\underline{a}}^{F^{-1}(q)} adF(a)k_2(x)}{dq^2} = -\frac{1}{2f(q)} < 0, \quad (\text{A12})$$

$$\frac{d^2 k_1(x)(1 - q) - \frac{1}{2} \int_{F^{-1}(q)}^{\bar{a}} adF(a)k_2(x)}{dq^2} = \frac{1}{2f(q)} k_2(x) > 0. \quad (\text{A13})$$

Combining $k_1(x) - \frac{1}{2} \int_{\underline{a}}^{F^{-1}(1)} adF(a)k_2(x) = k_1(x) - \frac{1}{2} \int_{\underline{a}}^{\bar{a}} adF(a)k_2(x) < 0$, (A8), (A10) and (A12), we have $\bar{q}(x) \in (0, 1)$. Combining $k_1(x)(1 - 0) - \frac{1}{2} \int_{F^{-1}(0)}^{\bar{a}} adF(a)k_2(x) = k_1(x) - \frac{1}{2} \int_{\underline{a}}^{\bar{a}} adF(a)k_2(x) < 0$, (A9), (A11) and (A13), we have $\underline{q}(x) = 0$.

Similarly, given $k_1(x) - \frac{1}{2} \int_{\underline{a}}^{\bar{a}} adF(a)k_2(x) > 0$, by (A8) - (A13), we have $\underline{q}(x) \in (0, 1)$ and $\bar{q}(x) = 1$.

A.6 Proof of Theorem 1

part 1 First consider the existence of the maximum point under condition 1.

Step 1: $[0, 1] \times \mathcal{Q}$ is compact.

A L -Lipschitz function u defined on an open interval (a, b) can be uniquely extended to a L -Lipschitz function on $[a, b]$. To see this, take a sequence $\{x_n\}_{n=1}^{\infty}$, $x_n \in (a, b)$ such that $\lim_{n \rightarrow \infty} x_n \rightarrow b$. By Lipschitzness of u , $\{u(x_n)\}_{n=1}^{\infty}$ is a Cauchy sequence, and thus $\lim_{n \rightarrow \infty} u(x_n)$ exists. The uniqueness of this limit among all such sequences then follows again from the Lipschitzness of u , i.e., $\lim_{x \rightarrow b} u(x)$ exists. Further, we have

$$|u(x_0) - u(b)| = \lim_{n \rightarrow \infty} |u(x_0) - u(x_n)| \leq \lim_{n \rightarrow \infty} L|x_0 - x_n| = L|x_0 - b|$$

for all $x_0 \in (a, b)$. The a side and the upper bound of $|u(a) - u(b)|$ is similar to the above.

Consider $\mathcal{Q}' = \left\{ q \in C_b([\underline{R}, \overline{R}]) : 0 \leq q(\cdot) \leq 1 \text{ and } \sup_{x \neq y} \frac{|q(x) - q(y)|}{|x - y|} \leq L \right\}$. For all $\varepsilon > 0$, choose $\delta = \varepsilon/L$. Then for all $q \in \mathcal{Q}'$, $x, y \in [\underline{R}, \overline{R}]$ and $|x - y| < \delta$ imply $|q(x) - q(y)| < \varepsilon$, i.e., $\mathcal{Q}'$ is uniformly equicontinuous. Since $\mathcal{Q}'$ is also uniformly bounded by definition, by Arzelà-Ascoli theorem, $\mathcal{Q}'$ is precompact in $C([\underline{R}, \overline{R}])$. By the unique extension, we can identify $\mathcal{Q}$ with $\mathcal{Q}'$, so $\mathcal{Q}$ is also precompact in $C_b([\underline{R}, \overline{R}])$. Let $\{q_n\}_{n=1}^\infty$ be a sequence in $\mathcal{Q}$ and $q_n \rightarrow q$ under the sup norm. For all $x, y \in (\underline{R}, \overline{R})$, we have

$$|q(x) - q(y)| = \lim_{n \rightarrow \infty} |q_n(x) - q_n(y)| \leq L|x - y|,$$

i.e., $\mathcal{Q}$ is closed and therefore compact. Then, $[0, 1] \times \mathcal{Q}$ which is the Cartesian product of two compact spaces is compact.

Step 2: The feasible set $D \cap ([0, 1] \times \mathcal{Q})$ is non-empty and compact.

By Step 1, we only need to verify that $D \cap ([0, 1] \times \mathcal{Q})$ is non-empty and closed. Note that $(x, 0) \in D$ for all $x \in [0, 1]$, which implies $D \cap ([0, 1] \times \mathcal{Q}) \neq \emptyset$. To verify closedness, let $\{(x_n, q_n)\}_{n=1}^\infty$ be a sequence in $D \cap ([0, 1] \times \mathcal{Q})$ such that $(x_n, q_n) \rightarrow (x, q)$ for $q \in \mathcal{Q}$. By definition, for all $x' \in [0, 1]$ and each $n \in \{1, 2, \dots\}$, we have

$$(Ax_n + \beta) \int_{\underline{R}}^{\overline{R}} q_n(x_n r_1) dG(r_1) \geq (Ax' + \beta) \int_{\underline{R}}^{\overline{R}} q_n(x' r_1) dG(r_1). \quad (\text{A14})$$

Since $|q_n(\cdot)| \leq 1$ and $q_n \rightarrow q$ uniformly, by dominated convergence theorem,

$$\int_{\underline{R}}^{\overline{R}} q_n(x' r_1) dG(r_1) \rightarrow \int_{\underline{R}}^{\overline{R}} q(x' r_1) dG(r_1), \quad \forall x' \in [0, 1].$$

On the other hand,

$$\begin{aligned} & \left| \int_{\underline{R}}^{\overline{R}} q_n(x_n r_1) dG(r_1) - \int_{\underline{R}}^{\overline{R}} q(x r_1) dG(r_1) \right| \\ & \leq \int_{\underline{R}}^{\overline{R}} |q_n(x_n r_1) - q(x r_1)| dG(r_1) \\ & \leq \int_{\underline{R}}^{\overline{R}} (|q_n(x_n r_1) - q_n(x r_1)| + |q_n(x r_1) - q(x r_1)|) dG(r_1) \\ & \leq \int_{\underline{R}}^{\overline{R}} \left(L|x_n - x| r_1 + \sup_{y \in (\underline{R}, \overline{R})} |q_n(y) - q(y)| \right) dG(r_1) \rightarrow 0. \end{aligned} \quad (\text{A15})$$

The convergence of the right hand side of the last inequality follows again from the dominated

convergence theorem. Therefore, as $n \rightarrow \infty$, the left hand side of (A14) converges and we get

$$(Ax + \beta) \int_{\underline{R}}^{\bar{R}} q(xr_1) dG(r_1) \geq (Ax' + \beta) \int_{\underline{R}}^{\bar{R}} q(x'r_1) dG(r_1).$$

By the continuity of $\bar{q}(\cdot)$, we also have $q(\cdot) = \lim_{n \rightarrow \infty} q_n(\cdot) \leq \lim_{n \rightarrow \infty} \bar{q}(x_n) = \bar{q}(x)$. Since a closed subset of a compact set is compact, $D \cap ([0, 1] \times \mathcal{Q})$ is compact.

Step 3: The objective function $O : D \cap ([0, 1] \times \mathcal{Q}) \rightarrow \mathbb{R}$ is continuous with respect to (x, q) .

By (A15) and the continuity of $k_1(x)$ and $k_2(x)$, we only need to consider the term

$$\int_{\underline{R}}^{\bar{R}} \int_{\underline{a}}^{F^{-1}(q(xr_1))} a dF(a) dG(r_1).$$

Let $\{(x_n, q_n)\}_{n=1}^{\infty}$ be a convergent sequence in $D \cap ([0, 1] \times \mathcal{Q})$ to (x, q) . Note that

$$\left| \int_{\underline{a}}^{F^{-1}(q_n(x_nr_1))} af(a) da - \int_{\underline{a}}^{F^{-1}(q(xr_1))} af(a) da \right| \leq K \left| F^{-1}(q_n(x_nr_1)) - F^{-1}(q(xr_1)) \right|$$

for some constant K . Since F is strictly increasing and continuous, F^{-1} is strictly increasing and continuous too. Similar to (A15), the right hand side of the above inequality converges to 0 as $n \rightarrow \infty$. Then by dominated convergence theorem, we have

$$\int_{\underline{R}}^{\bar{R}} \left(\int_{\underline{a}}^{F^{-1}(q_n(x_nr_1))} af(a) da \right) dG(r_1) \rightarrow \int_{\underline{R}}^{\bar{R}} \left(\int_{\underline{a}}^{F^{-1}(q(xr_1))} af(a) da \right) dG(r_1).$$

Here we use the obvious fact that $\left| \int_{\underline{a}}^{F^{-1}(q_n(x_nr_1))} af(a) da \right| \leq \mathbb{E}[a] < \infty$.

Now we get that the objective function $O(\cdot, \cdot)$ is continuous and the feasible set $D \cap ([0, 1] \times \mathcal{Q})$ is compact. By Weierstrass theorem (on extreme value), we know that there exist $(x^*, q^*) \in D \cap ([0, 1] \times \mathcal{Q})$ such that $O(x^*, q^*) = \sup_{(x, q) \in D \cap ([0, 1] \times \mathcal{Q})} O(x, q)$.

part 2 Generalizing the condition 1. to the condition 2. will not result in much change.

Step 1: $[0, 1] \times \mathcal{Q}$ is compact in the weak topology.

Recall that the condition 2. says that

$$\mathcal{Q} = \left\{ q \in W^{1,p}((\underline{R}, \bar{R})) : 1 < p < \infty, 0 \leq q(\cdot) \leq 1 \text{ and } \|Dq\|_p \leq L \right\}.$$

For all $q \in \mathcal{Q}$, $\|q\|_{1,p} = \|q\|_p + \|Dq\|_p \leq (|\underline{R}| + |\bar{R}|) + L$, i.e., $\mathcal{Q}$ is a bounded in norm. By Banach-Alaoglu-Bourbaki theorem and the fact that $W^{1,p}((\underline{R}, \bar{R}))$ is reflexive for $1 < p < \infty$ (see for example Theorem 3.6 of Adams and Fournier (2003)), we only need to verify that $\mathcal{Q}$ is closed in weak topology. Note that $\mathcal{Q}$ is convex, by Mazur lemma, the closedness of $\mathcal{Q}$ in weak topology

is equivalent to its closedness in strong topology, i.e., under the $\|\cdot\|_{1,p}$ norm. Let $\{q_n\}_{n=1}^\infty$ be a sequence in $\mathcal{Q}$ such that $\|q_n - q\|_{1,p} \rightarrow 0$ as $n \rightarrow \infty$. Since $\|\cdot\|_{1,p}$ dominates $\|\cdot\|_p$, we have $0 \leq q(\cdot) \leq 1$ a.e. with respect to Lebesgue measure. We also have that for all $\varepsilon > 0$, $\|Dq\|_p \leq \|Dq_n\|_p + \|Dq_n - Dq\|_p \leq \|Dq_n\|_p + \|q_n - q\|_{1,p} \leq L + \varepsilon$, when $n(\varepsilon)$ is chosen large enough. By the arbitrariness of ε , $\|Dq\| \leq L$, i.e., $q \in \mathcal{Q}$. Since the strong and weak topology is the same on $\mathbb{R}$, $[0, 1] \times \mathcal{Q}$ is compact in the weak topology (dual of product is isometric isomorphic to the product of the duals).

Step 2: The feasible set $D \cap ([0, 1] \times \mathcal{Q})$ is non-empty and weakly sequentially compact.

Since $0 \in W^{1,p}$, we already know that $D \cap ([0, 1] \times \mathcal{Q})$ is non-empty in the Step 2 of **part 1**. We only need to verify that $D \cap ([0, 1] \times \mathcal{Q})$ is weakly sequentially compact. Let $\{(x_n, q_n)\}_{n=1}^\infty$ be a sequence in $D \cap ([0, 1] \times \mathcal{Q})$. By definition, $\{(x_n, q_n)\}_{n=1}^\infty$ is bounded. By a version of Rellich-Kondrachov theorem (see for example Theorem 6.3 of [Adams and Fournier \(2003\)](#)), the embedding $W^{1,p} \hookrightarrow C^{0,1-\frac{1}{p}-\epsilon}$ is compact, for all $1 < p < \infty$ and for all $0 < \epsilon < 1 - \frac{1}{p}$. Hereafter, we always identify $q \in W^{1,p}(\underline{R}, \bar{R})$ with its $C^{0,1-\frac{1}{p}}(\underline{R}, \bar{R})$ version. This identification is possible due to Morrey inequality and the regularity of the boundary of $(\underline{R}, \bar{R})$, see for example Theorem 5 in subsection 5.8.4 of [Evans \(2010\)](#). To be more specific¹⁸, we show that $\{\underline{R}, \bar{R}\}$ is a C^1 (and thus Lipschitz) boundary. Consider $\{B(\underline{R}, \rho), B(\bar{R}, \rho)\}$, where $B(x_0, r)$ means the open ball centered at x_0 with radius r , and we may take $\rho = \frac{1}{4}(|\underline{R}| + |\bar{R}|)$. Let $\mathbb{R}^0 = \{0\}$, then $f_1 \equiv \underline{R}$, $f_2 \equiv \bar{R}$ are C^1 functions on $\mathbb{R}^0$. Obviously, $(\underline{R}, \bar{R}) \cap B(\underline{R}, \rho) = \{r \in B(\underline{R}, \rho) : r > f_1\}$ and $(\underline{R}, \bar{R}) \cap B(\bar{R}, \rho) = \{r \in B(\bar{R}, \rho) : r < f_2\}$, i.e., $\{\underline{R}, \bar{R}\}$ is the C^1 boundary of $(\underline{R}, \bar{R})$.

Now, by the compact embedding $W^{1,p} \hookrightarrow C^{0,1-\frac{1}{p}-\epsilon}$ and the discussion in the next step, we can modify (A15) to

$$\begin{aligned} & \left| \int_{\underline{R}}^{\bar{R}} q_n(x_n r_1) dG(r_1) - \int_{\underline{R}}^{\bar{R}} q(xr_1) dG(r_1) \right| \\ & \leq \int_{\underline{R}}^{\bar{R}} (|q_n(x_n r_1) - q_n(xr_1)| + |q_n(xr_1) - q(xr_1)|) dG(r_1) \\ & \leq \int_{\underline{R}}^{\bar{R}} \left(L'(|x_n r_1 - xr_1|)^{1-\frac{1}{p}-\epsilon} + \sup_{y \in (\underline{R}, \bar{R})} |q_n(y) - q(y)| \right) dG(r_1) \rightarrow 0, \end{aligned} \quad (\text{A16})$$

where L' is a constant, and q is a weak limit of $\{(x_n, q_n)\}_{n=1}^\infty$, extracting a subsequence if necessary. The existence of $q \in \mathcal{Q}$ is guaranteed by Eberlein-Šmulian theorem. The continuity of $\bar{q}(x)$ again guarantees that $q(\cdot) \leq \bar{q}(x)$.

Step 3: The objective function $O : D \cap ([0, 1] \times \mathcal{Q}) \rightarrow \mathbb{R}$ is weakly sequentially continuous, i.e., $(x_n, q_n) \rightarrow (x, q)$ implies that $O(x_n, q_n) \rightarrow O(x, q)$.

Let $\{(x_n, q_n)\}_{n=1}^\infty$ be a sequence in $D \cap ([0, 1] \times \mathcal{Q})$ such that $(x_n, q_n) \rightharpoonup (x, q)$ (weakly converges

¹⁸Although the boundary of $(\underline{R}, \bar{R})$ is simply $\{\underline{R}, \bar{R}\}$. Verifying the smooth boundary conditions can be quite confusing. For this reason, we belabor it a little bit here.

to (x, q)). By the discussions in the *Step 2* of **part 2**, $\{q_n\}$ converges in $C^{0,1-\frac{1}{p}-\epsilon}$ to some q' up to extraction of a subsequence. q and q' must coincide, otherwise $\int \mathbb{1}\{q > q'\}f =: \mathcal{L}_1(f) \neq 0$ or $\int \mathbb{1}\{q < q'\}f =: \mathcal{L}_2(f) \neq 0$. $\mathcal{L}_1$ and $\mathcal{L}_2$ are continuous linear functionals on $C^{0,1-\frac{1}{p}-\epsilon}$, but $\lim_{k \rightarrow \infty} \mathcal{L}_1(q_{n_k}) = \mathcal{L}_1(q') \neq \mathcal{L}_1(q)$ or $\lim_{k \rightarrow \infty} \mathcal{L}_2(q_{n_k}) = \mathcal{L}_2(q') \neq \mathcal{L}_2(q)$, a contradiction. Therefore $q = q'$, i.e., $q_n \rightarrow q$ in $C^{0,1-\frac{1}{p}-\epsilon}$. Now by (A16) and a discussions similar to *Step 3* of **part 1**, we can get the weak sequentially continuity of the objective function O .

Taking a maximizing sequence $\{(x_n, q_n)\}_{n=1}^\infty$ of the optimization problem, by *Step 2* of **part 2**, we can extract a weakly convergent subsequence $\{(x_{n_k}, q_{n_k})\}_{k=1}^\infty$. By *Step 3* of **part 2**, we have

$$O(x, q) = \lim_{k \rightarrow \infty} O(x_{n_k}, q_{n_k}) = \sup_{(x', q') \in D \cap ([0,1] \times \mathcal{Q})} O(x', q').$$

A.7 Proof of Lemma 3

Consider the auxiliary function function

$$I(q; x) := k_1(x)q - \frac{k_2(x)}{2} \int_{\underline{a}}^{F^{-1}(q)} adF(a), \quad q \in [0, 1].$$

By direct calculation, we have

$$\frac{dI^2(q; x)}{dx^2} = -\frac{k_2(x)}{2f(F^{-1}(q))} < 0, \quad \forall q \in [0, 1]. \quad (\text{A17})$$

The objective function $O(x, q)$ can be rewritten as $O(x, q) = \mathbb{E}[I(q(xr_1); x)]$. By (A17), $I(\cdot; x)$ is strictly concave for all $x \in [0, 1]$, so for q_1, q_2 such that $q_1(r) \neq q_2(r)$ on a set of positive Lebesgue measure, we have

$$\begin{aligned} O(x, \lambda q_1 + (1 - \lambda) q_2) &= \mathbb{E}[I(\lambda q_1(xr_1) + (1 - \lambda) q_2(xr_1))] \\ &> \mathbb{E}[\lambda I(q_1(xr_1)) + (1 - \lambda) I(q_2(xr_1))], \quad \forall x \in (0, 1], \forall \lambda \in (0, 1), \end{aligned}$$

since by assumption $g > 0$.

A.8 Proof of Proposition 5

Step 1: The existence of a maximizer.

The continuity of the objective function $O(q; x)$, and the compactness and closeness of $\mathcal{Q}$ are shown by the proof of Theorem 1. Therefore, we need to show $D_x \cap \mathcal{Q}$ is compact.

Let $\{q_n\}_{n \in \mathbb{N}}$ be a sequence in $D_x \cap \mathcal{Q}$ such that $q_n \rightarrow q$. By definition, we have

$$\begin{aligned} \lim_{n \rightarrow \infty} A \int_{\underline{R}}^{\bar{R}} q_n(xr_1) dG(r_1) + (Ax + \beta) \int_{\underline{R}}^{\bar{R}} q'_n(xr_1) r_1 dG(r_1) &= 0 \\ \Rightarrow A \int_{\underline{R}}^{\bar{R}} q(xr_1) dG(r_1) + (Ax + \beta) \int_{\underline{R}}^{\bar{R}} q'(xr_1) r_1 dG(r_1) &= 0. \end{aligned}$$

Since $\mathcal{Q}$ is compact and $D_x \cap \mathcal{Q}$ is closed, then $D_x \cap \mathcal{Q}$ is compact.

Step 2: The uniqueness of the maximizer. Denote the integrand of objective function $O(q; x)$ by $u_I(q(xr_1))$. The second derivative of $u_I(q(xr_1))$ with respect to $q(xr_1)$ is

$$-\left[\frac{1}{2} \frac{1}{f(F^{-1}(q(xr_1)))} k_2(x) \right] g(r_1) < 0, \text{ since } f(\cdot) \text{ and } g(\cdot) > 0.$$

Therefore $u_I(q(xr_1))$ is a strictly concave function with respect to $q(xr_1)$. For all $q_1, q_2 \in D_x \cap \mathcal{Q}$, $q_1(xr_1) \neq q_2(xr_1)$ for some xr_1 , we have

$$u_I[\lambda q_1(xr_1) + (1 - \lambda)q_2(xr_1)] > \lambda u_I[q_1(xr_1)] + (1 - \lambda)u_I[q_2(xr_1)]$$

for some xr_1 , where $\lambda \in (0, 1)$. It implies that

$$\begin{aligned} \int_{\underline{R}}^{\bar{R}} u_I[\lambda q_1(xr_1) + (1 - \lambda)q_2(xr_1)] dG(r_1) &> \int_{\underline{R}}^{\bar{R}} \lambda u_I[q_1(xr_1)] + (1 - \lambda)u_I[q_2(xr_1)] dG(r_1) \\ \Rightarrow \int_{\underline{R}}^{\bar{R}} u_I[(\lambda q_1 + (1 - \lambda)q_2)(xr_1)] dG(r_1) &> \lambda \int_{\underline{R}}^{\bar{R}} u_I[q_1(xr_1)] dG(r_1) + (1 - \lambda) \int_{\underline{R}}^{\bar{R}} u_I[q_2(xr_1)] dG(r_1). \end{aligned}$$

Here we use $D_x \cap \mathcal{Q}$ is a convex set. It implies that $O(q; x)$ is strictly convex with respect to q and the maximizer is unique.

A.9 Proof of Theorem 2

The proof is mainly following Theorem 2 in [Evans \(2010, p491\)](#).

Let $q^* \in D_x \cap \mathcal{Q}$ be the unique solution of the objective function. Fix any element $q_1 \in \mathcal{Q}$. Choose then any function $q_2 \in \mathcal{Q}$ with

$$\begin{aligned} &\frac{\int_{\underline{R}}^{\bar{R}} \frac{\partial I}{\partial y}(r_1, q^*, q^{*'})(q_1 - q^*) + \frac{\partial I}{\partial z}(r_1, q^*, q^{*'})(q'_1 - q^{*'}) dr_1}{\int_{\underline{R}}^{\bar{R}} \frac{\partial I}{\partial y}(r_1, q^*, q^{*'})(q_2 - q^*) + \frac{\partial I}{\partial z}(r_1, q^*, q^{*'})(q'_2 - q^{*'}) dr_1} < 0 \\ \Leftrightarrow &\frac{\int_{\underline{R}}^{\bar{R}} \left[\frac{\partial I}{\partial y}(r_1, q^*, q^{*'}) - \frac{d}{dr_1} \left(\frac{\partial I}{\partial z}(r_1, q^*, q^{*'}) \right) \right] (q_1 - q^*) dr_1 + (Ax + \beta) r_1 (q_1 - q^*) \Big|_{\underline{R}}^{\bar{R}}}{\int_{\underline{R}}^{\bar{R}} \left[\frac{\partial I}{\partial y}(r_1, q^*, q^{*'}) - \frac{d}{dr_1} \left(\frac{\partial I}{\partial z}(r_1, q^*, q^{*'}) \right) \right] (q_2 - q^*) dr_1 + (Ax + \beta) r_1 (q_2 - q^*) \Big|_{\underline{R}}^{\bar{R}}} < 0 \end{aligned}$$

and

$$\int_{\underline{R}}^{\overline{R}} \left[\frac{\partial I}{\partial y}(r_1, q^*, q^{*'}) - \frac{d}{dr_1} \left(\frac{\partial I}{\partial z}(r_1, q^*, q^{*'}) \right) \right] (q_2 - q^*) dr_1 + (Ax + \beta) r_1 (q_2 - q^*) \Big|_{\underline{R}}^{\overline{R}} \neq 0. \quad (\text{A18})$$

(A18) is possible because of Assumption 3. Then for each $0 \leq \tau \leq 1$ and $0 \leq \delta \leq 1$,

$$\begin{aligned} \tilde{q} &:= q^* + \tau(q_2 - q^*) + \delta[q_1 - (q^* + \tau(q_2 - q^*))] \in \mathcal{Q} \\ &\Leftrightarrow (1 - \delta)(1 - \tau)q^* + (1 - \delta)\tau q_2 + \delta q_1 \in \mathcal{Q} \end{aligned}$$

since $\mathcal{Q}$ is convex. Now write

$$i(\tau, \delta) := \int_{\underline{R}}^{\overline{R}} I(r_1, (1 - \delta)(1 - \tau)q^* + (1 - \delta)\tau q_2 + \delta q_1, (1 - \delta)(1 - \tau)q^{*'} + (1 - \delta)\tau q_2' + \delta q_1') dr_1.$$

Clearly,

$$i(0, 0) = \int_{\underline{R}}^{\overline{R}} I(r_1, q^*, q^{*'}) dr_1 = 0.$$

In addition, i is C^1 and

$$\frac{\partial i}{\partial \tau}(\tau, \delta) = (1 - \delta) \int_{\underline{R}}^{\overline{R}} \frac{\partial I}{\partial y}(r_1, \tilde{q}, \tilde{q}')(q_2 - q^*) + \frac{\partial I}{\partial z}(r_1, \tilde{q}, \tilde{q}')(q_2' - q^{*'}) dr_1, \quad (\text{A19})$$

$$\begin{aligned} \frac{\partial i}{\partial \delta}(\tau, \delta) &= \int_{\underline{R}}^{\overline{R}} \frac{\partial I}{\partial y}(r_1, \tilde{q}, \tilde{q}')[q_1 - (q^* + \tau(q_2 - q^*))] \\ &\quad + \frac{\partial I}{\partial z}(r_1, \tilde{q}, \tilde{q}')[q_1' - (q^{*'} + \tau(q_2' - q^{*'}))] dr_1. \end{aligned} \quad (\text{A20})$$

Consequently (A18) implies that

$$\frac{\partial i}{\partial \tau}(0, 0) \neq 0.$$

According to the Implicit Function Theorem, there exists a C^1 function $\phi : R \rightarrow R$ such that $\phi(0) = 0$ and

$$i(\phi(\delta), \delta) = 0 \quad (\text{A21})$$

for all sufficiently small δ . Differentiating, we discover that

$$\frac{\partial i}{\partial \tau}(\phi(\delta), \delta)\phi'(\delta) + \frac{\partial i}{\partial \delta}(\phi(\delta), \delta) = 0,$$

whence (A19) and (A20) yield

$$\phi'(0) = -\frac{\frac{\partial i}{\partial \delta}(\phi(0), 0)}{\frac{\partial i}{\partial \tau}(\phi(0), 0)} = -\frac{\int_{\underline{R}}^{\bar{R}} \frac{\partial I}{\partial y}(r_1, q^*, q^{*'}) (q_1 - q^*) + \frac{\partial I}{\partial z}(r_1, q^*, q^{*'}) (q'_1 - q^{*'}) dr_1}{(1 - \delta) \int_{\underline{R}}^{\bar{R}} \frac{\partial I}{\partial y}(r_1, q^*, q^{*'}) (q_2 - q^*) + \frac{\partial I}{\partial z}(r_1, q^*, q^{*'}) (q'_2 - q^{*'}) dr_1}. \quad (\text{A22})$$

Under Assumption 4, we have $\phi'(0) > 0$. By $\phi(0) = 0$, $\phi'(0) > 0$ implies that $0 \leq \phi(\delta) \leq 1$ for sufficiently small positive δ , say $0 \leq \delta \leq \delta_0$.

Now set

$$\tilde{q}(\delta) := \phi(\delta)(q_2 - q^*) + \delta[q_1 - (q^* + \phi(\delta)(q_2 - q^*))]$$

and write

$$\begin{aligned} o(\delta) &:= O(q^* + \tilde{q}(\delta); x) \\ &= \int_{\underline{R}}^{\bar{R}} k_1(x)(q^* + \tilde{q}(\delta))(xr_1) - \frac{1}{2}k_2(x) \left(\int_{\underline{a}}^{F^{-1}((q^* + \tilde{q}(\delta))(xr_1))} a dF(a) \right) dG(r_1). \end{aligned}$$

Since (A21), we see that $q^* + \tilde{q}(\delta) \in D_x \cap \mathcal{Q}$.

By q^* maximizes $O(q; x)$, we have $o(0) \geq o(\delta)$ for all $0 \leq \delta \leq 1$. Hence

$$0 \geq o'(0) = \int_{\underline{R}}^{\bar{R}} k_1(x) [\phi'(0)(q_2 - q^*) + q_1 - q^*] - \frac{1}{2}k_2(x) F^{-1}(q^*) [\phi'(0)(q_2 - q^*) + q_1 - q^*] dG(r_1). \quad (\text{A23})$$

Define

$$\lambda := \frac{\int_{\underline{R}}^{\bar{R}} k_1(x)(q_2 - q^*) - \frac{1}{2}k_2(x) F^{-1}(q^*)(q_2 - q^*) dG(r_1)}{(1 - \delta) \int_{\underline{R}}^{\bar{R}} \frac{\partial I}{\partial y}(r_1, q^*, q^{*'}) (q_2 - q^*) + \frac{\partial I}{\partial z}(r_1, q^*, q^{*'}) (q'_2 - q^{*'}) dr_1}, \quad (\text{A24})$$

then we see that

$$\begin{aligned} \phi'(0) &\int_{\underline{R}}^{\bar{R}} k_1(x)(q_2 - q^*) - \frac{1}{2}k_2(x) F^{-1}(q^*)(q_2 - q^*) dG(r_1) \\ &= -\lambda \int_{\underline{R}}^{\bar{R}} \frac{\partial I}{\partial y}(r_1, q^*, q^{*'}) (q_1 - q^*) + \frac{\partial I}{\partial z}(r_1, q^*, q^{*'}) (q'_1 - q^{*'}) dr_1 \end{aligned}$$

and plugging this into (A23), we have

$$0 \geq \int_{\underline{R}}^{\bar{R}} (q_1 - q^*) \left[k_1(x) - \frac{1}{2}k_2(x) F^{-1}(q^*) \right] dG(r_1)$$

$$\begin{aligned}
 & -\lambda \int_{\underline{R}}^{\bar{R}} \frac{\partial I}{\partial y}(r_1, q^*, q^{*'})(q_1 - q^*) + \frac{\partial I}{\partial z}(r_1, q^*, q^{*'})(q'_1 - q^{*'}) dr_1 \\
 &= \int_{\underline{R}}^{\bar{R}} (q_1 - q^*) \left[k_1(x) - \frac{1}{2} k_2(x) F^{-1}(q^*) \right] dG(r_1) - \lambda \int_{\underline{R}}^{\bar{R}} A(q_1 - q^*) + (Ax + \beta) r_1 (q'_1 - q^{*'}) dG(r_1) \\
 &= \int_{\underline{R}}^{\bar{R}} (q_1 - q^*) \left[k_1(x) - \frac{1}{2} k_2(x) F^{-1}(q^*) \right] dG(r_1) - \lambda [Ag(r_1) - (Ax + \beta)(g(r_1) + r_1 g'(r_1))/x] dr_1 \\
 &\quad + \lambda (Ax + \beta) r_1 g(r_1)/x [q_1(xr_1) - q^*(xr_1)] \Big|_{\underline{R}}^{\bar{R}} \\
 &= \int_{\underline{R}}^{\bar{R}} (q_1 - q^*) \left[k_1(x) - \frac{1}{2} k_2(x) F^{-1}(q^*) + \lambda \beta/x + \lambda (A + \beta/x) r_1 g'(r_1)/g(r_1) \right] dG(r_1) \\
 &\quad + \lambda (Ax + \beta) r_1 g(r_1)/x [q_1(xr_1) - q^*(xr_1)] \Big|_{\underline{R}}^{\bar{R}}.
 \end{aligned}$$

Example 1. Consider that the risk return R_t follows a uniform distribution $U[\underline{R}, \bar{R}]$, indicating that $r_1 g'(r_1) + g(r_1)$ is constant and thus the Euler-Lagrange approach under Assumption 3 is not appropriate. Let $q(\cdot)$ simply takes a bang-bang form:

$$q(r; x, \hat{q}) = \begin{cases} \hat{q}, & R \in [x\bar{R}, x\underline{R}]; \\ 0, & \text{otherwise,} \end{cases}$$

where $x \in [0, 1]$. Under the bang-bang probability $q(r; x^*, \hat{q})$, the incentive constraint (6) is

$$\begin{aligned}
 & x \in \arg \max_{x'} \left\{ (x' A + \beta) \int_{\min\{x^*, x'\}\underline{R}}^{\min\{x^*, x'\}\bar{R}} \hat{q} dG(r/x') \right\} \\
 & \Rightarrow x \in \arg \max_{x'} \left\{ (x' A + \beta) \hat{q} \left[G(\min\{x^*, x'\}\bar{R}/x') - G(\min\{x^*, x'\}\underline{R}/x') \right] \right\}.
 \end{aligned}$$

Note that

$$\begin{aligned}
 & \arg \max_{x' \leq x^*} (x' A + \beta) \hat{q} \left[G(x'\bar{R}/x') - G(x'\underline{R}/x') \right] = x^*; \\
 & \arg \max_{x' \geq x^*} (x' A + \beta) \hat{q} \left[G(x^*\bar{R}/x') - G(x^*\underline{R}/x') \right] = \arg \max_{x' \geq x^*} x^* (A + \beta/x') \hat{q} = x^*,
 \end{aligned}$$

implying that x^* is always incentive compatible without information rent. This result reflects the fact that algorithms do influence managers' decisions and rectify contractual flaws. In particular, when returns are uniformly distributed, this algorithm achieves exactly the first best equilibrium.

A.10 Proof of Proposition 6

Given x^* and $\lambda \neq 0$, we have

$$\mathbb{E}[R_2 - \phi(R_2)|x^*] - \frac{1}{2}F^{-1}\left(\frac{k_1(x^*)}{k_2(x^*)/2}\right)\mathbb{E}[(R_2 - \phi(R_2))^2|x^*] = 0.$$

Plug $\lambda \neq 0$ into (20), there exists $r, r' \in [\underline{R}, \bar{R}]$ such that

$$\hat{a}^*(r) = F^{-1}(q^*(r)) > F^{-1}\left(\frac{k_1(x^*)}{k_2(x^*)/2}\right) > \hat{a}^*(r') = F^{-1}(q^*(r')).$$

Then we obtain

$$\mathbb{E}[R_2 - \phi(R_2)|x^*] - \frac{1}{2}\hat{a}^*(r)\mathbb{E}[(R_2 - \phi(R_2))^2|x^*] < 0 < \mathbb{E}[R_2 - \phi(R_2)|x^*] - \frac{1}{2}\hat{a}^*(r')\mathbb{E}[(R_2 - \phi(R_2))^2|x^*].$$

B Alternative Timeline and Information Settings

In the baseline model, the roles of the recommendation algorithm are twofold. (i) It processes two information, fund historical returns and investors' risk aversions. (ii) By algorithmic automation, it provides a commitment power which ensures execution even realizes cases with ex-post inefficiency. To further understand the economic meaning of using an algorithm, we separately mute the above functionalities by considering alternative information structures and timeline, as shown in Figure 3, then analyze what the equilibrium would be. For tractability, we follow the discrete setting in Section 4. We analyze each case respectively, and end up with a visualization of their comparison in Figure 4.

Blind investment. In a primordial case, non-professional investors do not adopt a platform, but meet the fund manager by chance. They do not know about the historical returns R_1 and their risk aversion a as Timeline 2 describes. They set up a belief about their risk aversions, say a population average risk aversion, and decide the investment choices.

Proposition 7. *Given the contract $\phi(\cdot)$, $x = 1$ is a dominant strategy of the fund manager.*

1. *If $k_1(1) - 1/2\mathbb{E}[a]k_2(1) \geq 0$, there exists an equilibrium where all investors invest in the fund and the manager chooses $x^* = 1$.*
2. *If $k_1(1) - 1/2\mathbb{E}[a]k_2(1) < 0$, there exists an equilibrium where no investor invest in the fund and the manager chooses $x^* = 1$.*

When the investor has no information, Proposition 7 describes that, given a realistic simple

contract, the trade can only be made under $x = 1$ according to the manager's risk-neutral preferences, regardless of the distribution of the investor's own risk aversion. Therefore, there is no risk sharing at this point. Managers chase on risks, and investors inappropriately take the risk.

Investment on fund distribution platforms. We consider a scenario (captured by Timeline 3) where investors see the fund on a distribution platform which provides the historical return R_1 . This setup is equivalent to assuming that the platform simply aggregates the historical returns of the fund, corresponding to the context of [Hong et al. \(2024\)](#). Also, they make decisions based on their belief, the population average risk aversion. Denote the strategy of investors by $q_I : [\underline{R}, \bar{R}] \rightarrow [0, 1]$.

Proposition 8. *Suppose the contract $\phi(\cdot)$ is given and the investors can observe the realization of historical return R_1 .*

1. *If $k_1(1) - 1/2\mathbb{E}[a]k_2(1) \geq 0$, there exists an equilibrium where all investors invest in the fund and the manager chooses $x^* = 1$.*
2. *If $k_1(1) - 1/2\mathbb{E}[a]k_2(1) < 0$, there is no equilibrium where the expected payoff of investors is strictly positive, but exists an equilibrium $(\bar{x}_I, q^*(\cdot))$,*

$$q^*(r) = \begin{cases} 0, & r \in [\underline{R}, \bar{x}_I \underline{R}] \cup (\underline{x}_I \underline{R}, 0) \cup (0, \underline{x}_I \bar{R}) \cup (\bar{x}_I \bar{R}, \bar{R}); \\ 1, & r \in [\bar{x}_I \underline{R}, \underline{x}_I \underline{R}] \cup [\underline{x}_I \bar{R}, \bar{x}_I \bar{R}]; \\ q(0), & r = 0, \end{cases}$$

where

$$\begin{aligned} \bar{x}_I &= \max\{x \in [0, 1] : k_1(x) - 1/2\mathbb{E}[a]k_2(x) \geq 0\}, \\ \underline{x}_I &= \min\{x \in [0, 1] : k_1(x) - 1/2\mathbb{E}[a]k_2(x) \geq 0\}, \\ q(0) &\in \left[0, \min\left\{\frac{(A\bar{x}_I + \beta)(1 - p(0))}{A(1 - \bar{x}_I)p(0)}, 1\right\}\right] \end{aligned}$$

and the expected payoff of investors is zero.

Compared to classic principal-agent problems, the platform empowers investors to form decisions in response to historical returns, which allows fund managers to credibly deliver (noisy at $r = 0$) signals about x , thus facilitating trading. Therefore, compared to blind investment, the platform can make the equilibrium with investment always exist.

Compared to the baseline model in [Figure 1](#), it is not enough to increase the investor's expected

return by allowing investors to observe historical returns. Because the fund manager takes actions first and is unable to adjust x later, the investor could always choose to buy once they infer an x (when the historical return is non-zero) that generates a positive expected return.¹⁹ Anticipating this, the fund manager will always increase x to x_I , where the investor's expected return is 0. The underlying reason is that any atomic investor does not have the bargaining power on x , and also cannot commit to invest when a proper x is observed.

To further see the power of commitment, we can further suppose the support of risky returns is $\{\underline{R}, \bar{R}\}$, i.e., investors always correctly learn x . Then we have the following implication.

Corollary 2. *Suppose $\text{supp}(R_t) = \{\underline{R}, \bar{R}\}$ and $k_1(1) - 1/2\mathbb{E}[a]k_2(1) < 0$, there is no equilibrium with strictly positive investor expected payoff, even though there is no information asymmetry.*

In addition, when the realized return is zero, investors cannot recognize x , similar to the baseline. This leads to conservative investments by investors to avoid the fund manager's deviation. Thus there is still a no-trade efficiency loss due to the non-informative signal $r = 0$.

Additional information about the fund. Continue with the previous case. A natural idea for the platform is to provide information about the fund in addition to an observation of the historical return, e.g., a series of historical returns, Sharpe ratios, etc. These cases can be directly analyzed within our continuous framework—the additional information solely contributes to the conditional probabilities. Align with practice, these achievements can increase the confidence of inferring the allocation. However, they still fail to replace the algorithm, as the algorithm also processes information about investors' risk aversion. Therefore, there is no fundamental change if just allowing for additional information about the fund.

Investment experts. Then does algorithm only serve as informing investors about their risk aversion? We consider the case (captured by Timeline 4) when the investors are experts: they observe the historical return R_1 and know the risk aversion a . Then their investment choice is determined by two information, somehow similar to the algorithm, denoted as $m_I : [a, \bar{a}] \times [\underline{R}, \bar{R}] \rightarrow [0, 1]$. For any a , denote the minimum and maximum of $\{x \in [0, 1] : k_1(x) - 1/2ak_2(x) \geq 0\}$ as $\underline{x}_I(a)$ and $\bar{x}_I(a)$, respectively, and $\hat{a}(x) = k_1(x)/(k_2(x)/2)$. Proposition 9 analytically solves the equilibrium allocation and the corresponding investment choice.

¹⁹When the distribution of risky returns is continuous, the investors have a noisy signal about x .

Proposition 9. For any equilibrium (x^*, m^*) (if exists), x^* satisfies that for any $x \in [0, 1]$,

$$\begin{aligned} (x^*A + \beta) & \left[p(0) \int_{\underline{a}}^{\hat{a}(x^*)} 1dF(a) + p(\bar{R}) \int_{\underline{a}}^{\hat{a}(x^*)} 1dF(a) + p(\underline{R}) \int_{\underline{a}}^{\hat{a}(x^*)} 1dF(a) \right] \\ & \geq (xA + \beta) \left[p(0) \int_{\underline{a}}^{\hat{a}(x)} 1dF(a) + p(\bar{R}) \int_{\underline{a}}^{\hat{a}(x)} 1dF(a) + p(\underline{R}) \int_{\underline{a}}^{\hat{a}(x)} 1dF(a) \right]. \end{aligned} \quad (\text{B25})$$

For $a < \hat{a}(x^*)$,

$$m_I^*(r, a) = \begin{cases} 0, & \text{if } r \in [\underline{R}, \bar{x}_I(a)\underline{R}] \cup (\underline{x}_I(a)\underline{R}, 0) \cup (0, \underline{x}_I(a)\bar{R}) \cup (\bar{x}_I(a)\bar{R}, \bar{R}), \\ 1, & \text{if } r \in [\bar{x}_I(a)\underline{R}, \underline{x}_I(a)\underline{R}] \cup [\underline{x}_I(a)\bar{R}, \bar{x}_I(a)\bar{R}] \cup \{0\}; \end{cases}$$

For $a > \hat{a}(x^*)$,

$$m_I^*(r, a) = \begin{cases} 0, & \text{if } r \in [\underline{R}, \bar{x}_I(a)\underline{R}] \cup (\underline{x}_I(a)\underline{R}, 0) \cup (0, \underline{x}_I(a)\bar{R}) \cup (\bar{x}_I(a)\bar{R}, \bar{R}) \cup \{0\}, \\ 1, & \text{if } r \in [\bar{x}_I(a)\underline{R}, \underline{x}_I(a)\underline{R}] \cup [\underline{x}_I(a)\bar{R}, \bar{x}_I(a)\bar{R}]; \end{cases}$$

For $a = \hat{a}(x^*)$,

$$m_I^*(r, a) = \begin{cases} 0, & \text{if } r \in [\underline{R}, \bar{x}_I(a)\underline{R}] \cup (\underline{x}_I(a)\underline{R}, 0) \cup (0, \underline{x}_I(a)\bar{R}) \cup (\bar{x}_I(a)\bar{R}, \bar{R}), \\ 1, & \text{if } r \in (\bar{x}_I(a)\underline{R}, \underline{x}_I(a)\underline{R}) \cup (\underline{x}_I(a)\bar{R}, \bar{x}_I(a)\bar{R}); \end{cases}$$

and $m_I^*(r, a)$ can be arbitrarily assigned in $[0, 1]$ when $r \in \{\bar{x}_I(a)\underline{R}, \underline{x}_I(a)\underline{R}, \underline{x}_I(a)\bar{R}, \bar{x}_I(a)\bar{R}, 0\}$.

Compared to Proposition 8, the additional information about a allows investors to always realize non-negative expected payoff, resulting in a strictly positive aggregate investor expected utility. The manager no longer uses $\mathbb{E}[a]$ in deciding x , but considers the entire distribution F . Therefore, when the historical return is informative, the manager does suffer a punishment of risk chasing, as high-risk-aversion investors would exit. On the other hand, investors with sufficiently low risk aversion always invest, even without any information about x . Therefore, investors cannot generate enough punishment when the historical return is not informative, making the resulting equilibrium allocation x^* higher than in the baseline case.

So far, we have seen the algorithm serves not only an information delivery mechanism, but also a source of commitment power. Even with experts, the algorithm can add noise to their risk aversion observations by a threshold function, thus realizing that fewer investors would like to invest when the historical return is less informative, thus alleviating fund managers' incentives to raise x .

B.1 Proof of Proposition 7

Obviously, $x = 1$ is a dominant strategy of the fund manager. Given $x = 1$ and the expected level of risk aversion is $\mathbb{E}[a]$, if the expected utility of investment is non-negative, $k_1(1) - 1/2\mathbb{E}[a]k_2(1) \geq 0$, investors could choose to invest and all agents do not deviate. Conversely, if the expected utility of investment is negative, that is $k_1(1) - 1/2\mathbb{E}[a]k_2(1) < 0$, the investors do not invest, and all agents do not deviate.

B.2 Proof of Proposition 8

If $k_1(1) - 1/2\mathbb{E}[a]k_2(1) \geq 0$, then investors should always invest in the fund given $x = 1$. Given $q_I(\cdot) \equiv 1$, the manager prefer $x = 1$.

If $k_1(1) - 1/2\mathbb{E}[a]k_2(1) < 0$, we define the minimum and maximum of $\{x \in [0, 1] : k_1(x) - 1/2\mathbb{E}[a]k_2(x) \geq 0\}$ as $\underline{x}_I$ and $\bar{x}_I \in (0, 1)$, respectively. Here we use the fact that $k_1(x) - 1/2\mathbb{E}[a]k_2(x)$ is concave and strictly continuous with respect to x . Since $k_1(0) - 1/2\mathbb{E}[a]k_2(0) < 0$, we have $\underline{x}_I > 0$. Consider the strategy of the investors. When the historical return is non-zero, investors can know the fund manager's choice of x . As a result, consider the subgames, given $r \in (\bar{x}_I \underline{R}, \underline{x}_I \underline{R}) \cup (\underline{x}_I \bar{R}, \bar{x}_I \bar{R})$, investors always should invest in the fund, that is $q_I(r) = 1$. Similarly, given $r \in [\underline{R}, \bar{x}_I \underline{R}) \cup (\underline{x}_I \underline{R}, 0) \cup (0, \underline{x}_I \bar{R}) \cup (\bar{x}_I \bar{R}, \bar{R})$, investors should not invest in the fund, that is $q_I(r) = 0$. As for $r \in \{\underline{x}_I \underline{R}, \underline{x}_I \bar{R}, \bar{x}_I \underline{R}, \bar{x}_I \bar{R}\}$, investors are indifferent to $q_I(r) = q_r \in [0, 1]$, because

$$q_r [k_1(x) - 1/2\mathbb{E}[a]k_2(x)] = q_r \times 0 = 0, \forall q_r \in [0, 1].$$

The investor cannot recognize x when the return is 0, so $q_I(0)$ depends on the equilibrium we consider.

Given $k_1(1) - 1/2\mathbb{E}[a]k_2(1) < 0$, we suppose there exists an equilibrium $(x^*, q_I^*(\cdot))$ where the investors invest in the fund with some strictly positive probability, which implies $x^* \in [\underline{x}_I, \bar{x}_I]$. Given the above response of investors, we know that the manager prefers $\sup[\underline{x}_I, \bar{x}_I] = \bar{x}_I$ to any $x \in [0, \bar{x}_I)$. This is because,

$$\begin{aligned} & (xA + \beta)[q_I(0)p(0) + q_I(\underline{x}\underline{R})p(\underline{R}) + q_I(\underline{x}\bar{R})p(\bar{R})] \\ &= (xA + \beta)[q_I(0)p(0) + p(\underline{R}) + p(\bar{R})] \\ &\leq \sup_{x \in (\underline{x}_I, \bar{x}_I)} (xA + \beta)[q_I(0)p(0) + p(\underline{R}) + p(\bar{R})], \text{ for all } x \in (\underline{x}_I, \bar{x}_I), \end{aligned}$$

and

$$(xA + \beta)[q_I(0)p(0) + q_I(\underline{x}\underline{R})p(\underline{R}) + q_I(\underline{x}\bar{R})p(\bar{R})]$$

$$\begin{aligned}
 &= (xA + \beta)q_I(0)p(0) \\
 &\leq \sup_{x \in (\underline{x}_I, \bar{x}_I)} (xA + \beta)[q_I(0)p(0) + p(\underline{R}) + p(\bar{R})], \text{ for all } x \in (0, \underline{x}_I),
 \end{aligned}$$

If $x^* \in [0, \bar{x}_I)$, the manager can always deviate from x^* to choose the larger $x' \in (x^*, \bar{x}_I)$. Therefore, if there exists such an equilibrium characterized by (x^*, q_I^*) , $x^* = \sup(\underline{x}_I, \bar{x}_I) = \bar{x}_I$, otherwise the manager always deviate x^* . This also means that the expected payoff of investors must be 0 according to the definition of $\bar{x}_I$. Then the first necessary condition of the equilibrium is that $q^*(\bar{x}_I \underline{R})$ and $q^*(\bar{x}_I \bar{R})$ need to satisfy

$$\begin{aligned}
 &(A\bar{x}_I + \beta)[q_I(0)p(0) + q_I(\bar{x}_I \underline{R})p(\underline{R}) + q_I(\bar{x}_I \bar{R})p(\bar{R})] \\
 &\geq \sup_{x \in [\underline{x}_I, \bar{x}_I]} \{Ax + \beta\}[q_I(0)p(0) + p(\underline{R}) + p(\bar{R})] = (A\bar{x}_I + \beta)q_I(0)p(0) \\
 &\Rightarrow q_I(\bar{x}_I \bar{R}) = q_I(\bar{x}_I \underline{R}) = 1,
 \end{aligned}$$

where the last line is because, if $q_I(\bar{x}_I \bar{R}) < 1$ or $q_I(\bar{x}_I \underline{R}) < 1$, there always exists $x' \in (\underline{x}_I, \bar{x}_I)$ such that

$$\begin{aligned}
 &(A\bar{x}_I + \beta)[q_I(0)p(0) + q_I(\bar{x}_I \underline{R})p(\underline{R}) + q_I(\bar{x}_I \bar{R})p(\bar{R})] \\
 &< (Ax' + \beta)[q_I(0)p(0) + p(\underline{R}) + p(\bar{R})] = (Ax' + \beta)q_I(0)p(0).
 \end{aligned}$$

At the same time, the second necessary condition is that, $q_I(0)$ need to satisfy

$$\begin{aligned}
 &(A\bar{x}_I + \beta)[q_I(0)p(0) + q_I(\bar{x}_I \underline{R})p(\underline{R}) + q_I(\bar{x}_I \bar{R})p(\bar{R})] \\
 &= (A\bar{x}_I + \beta)[q_I(0)p(0) + p(\underline{R}) + p(\bar{R})] \\
 &\geq \sup_{x \in (\bar{x}_I, 1]} (Ax' + \beta)[q_I(0)p(0) + q_I(x' \underline{R})p(\underline{R}) + q_I(x' \bar{R})p(\bar{R})] \\
 &= (A + \beta)q_I(0)p(0) \\
 &\Leftrightarrow (A\bar{x}_I + \beta)[q_I(0)p(0) + p(\underline{R}) + p(\bar{R})] \geq (A + \beta)q_I(0)p(0) \\
 &\Leftrightarrow q(0) \leq \frac{(A\bar{x}_I + \beta)(1 - p(0))}{A(1 - \bar{x}_I)p(0)}.
 \end{aligned}$$

The second line uses $q_I(\bar{x}_I \bar{R}) = q_I(\bar{x}_I \underline{R}) = 1$ in the equilibrium. The forth line uses $q_I(x' \bar{R}) = q_I(x' \underline{R}) = 0$ for all $x' \in (\bar{x}_I, 1]$.

In addition, we can see that $(\bar{x}_I, q_I^*(\cdot))$ is a PBE, where

$$q_I^*(r) = \begin{cases} 0 & , \text{ if } r \in [\underline{R}, \bar{x}_I \underline{R}] \cup (\underline{x}_I \underline{R}, 0) \cup (0, \underline{x}_I \bar{R}) \cup (\bar{x}_I \bar{R}, \bar{R}) \\ 1 & , \text{ if } r \in [\bar{x}_I \underline{R}, \underline{x}_I \underline{R}] \cup [\underline{x}_I \bar{R}, \bar{x}_I \bar{R}] \\ q(0) & , \text{ if } r = 0 \end{cases}$$

and

$$q(0) \in \left[0, \min \left\{ \frac{(A\bar{x}_I + \beta)(1 - p(0))}{A(1 - \bar{x}_I)p(0)}, 1 \right\} \right].$$

B.3 Proof of Proposition 9

Consider the subgames with given historical r , for any a , if $r \in (\bar{x}_I(a) \underline{R}, \underline{x}_I(a) \underline{R}) \cup (\underline{x}_I(a) \bar{R}, \bar{x}_I(a) \bar{R})$, investors with a always should invest in the fund, that is $m(r, a) = 1$. Similarly, given $r \in [\underline{R}, \bar{x}_I(a) \underline{R}] \cup (\underline{x}_I(a) \underline{R}, 0) \cup (0, \underline{x}_I(a) \bar{R}) \cup (\bar{x}_I(a) \bar{R}, \bar{R})$, investors should not invest in the fund, that is $m(r, a) = 0$. As for $r \in \{\underline{x}_I(a) \underline{R}, \underline{x}_I(a) \bar{R}, \bar{x}_I(a) \underline{R}, \bar{x}_I(a) \bar{R}\}$, investors are indifferent to $m(r, a) = m_{r,a} \in [0, 1]$.

Note that $k_1(x) - 1/2ak_2(x)$ is strictly decreasing with respect to a . Then we know that, when $\underline{x}_I(a) \in (0, 1)$ and $\bar{x}_I(a) \in (0, 1)$, $\underline{x}_I(a)$ strictly decreases with a and $\bar{x}_I(a)$ strictly increases with a . It implies that if $m(\bar{R}x, a) = 1$ for some a , then $m(\bar{R}x, a') = 1$ for all $a' \leq a$.

Define $\hat{a}(x)$ as $\sup\{a \in [\underline{a}, \bar{a}] : m(x\bar{R}, a) = 1\} = k_1(x)/(k_2(x)/2)$.

Then, the expected payoff of the manager is

$$\begin{aligned} & (xA + \beta) \left[p(0) \int_{\underline{a}}^{\bar{a}} m(a, 0) dF(a) + p(\bar{R}) \int_{\underline{a}}^{\bar{a}} m(a, x\bar{R}) dF(a) + p(\underline{R}) \int_{\underline{a}}^{\bar{a}} m(a, x\underline{R}) dF(a) \right] \\ &= (xA + \beta) \left[p(0) \int_{\underline{a}}^{\bar{a}} m(a, 0) dF(a) + p(\bar{R}) \int_{\underline{a}}^{\hat{a}(x)} 1 dF(a) + p(\underline{R}) \int_{\underline{a}}^{\hat{a}(x)} 1 dF(a) \right. \\ & \quad \left. + p(\bar{R}) \int_{\hat{a}(x)}^{\bar{a}} 0 dF(a) + p(\underline{R}) \int_{\hat{a}(x)}^{\bar{a}} 0 dF(a) \right] \\ &= (xA + \beta) \left[p(0) \int_{\underline{a}}^{\bar{a}} m(a, 0) dF(a) + p(\bar{R}) \int_{\underline{a}}^{\hat{a}(x)} 1 dF(a) + p(\underline{R}) \int_{\underline{a}}^{\hat{a}(x)} 1 dF(a) \right]. \end{aligned}$$

In the second line, we use the fact that the value of $m(\hat{a}, x\bar{R})$ and $m(\hat{a}, x\underline{R})$ does not influence the value of the integral.

Then we suppose there exists an equilibrium (x^*, m_I^*) . Different with Proposition 8, on the equilibrium path, investors observing $a < \hat{a}(x^*)$ invest in the fund even if the historical return is 0, that is $m(a, 0) = 1$. Meanwhile, on the equilibrium path, investors observing $a > \hat{a}(x^*)$ do not

invest in the fund if the historical return is 0, that is $m(a, 0) = 0$. Given the above response of investors, one of the necessary conditions of equilibrium (x^*, m_l^*) is that the manager prefers x^* to any $x \in [0, x^*)$

$$\begin{aligned} (x^*A + \beta) & \left[p(0) \int_{\underline{a}}^{\hat{a}(x^*)} 1dF(a) + p(\bar{R}) \int_{\underline{a}}^{\hat{a}(x^*)} 1dF(a) + p(\underline{R}) \int_{\underline{a}}^{\hat{a}(x^*)} 1dF(a) \right] \\ & \geq (xA + \beta) \left[p(0) \int_{\underline{a}}^{\hat{a}(x)} 1dF(a) + p(\bar{R}) \int_{\underline{a}}^{\hat{a}(x)} 1dF(a) + p(\underline{R}) \int_{\underline{a}}^{\hat{a}(x)} 1dF(a) \right]. \end{aligned}$$

If the manager chooses a lower x , it can increase the probability of being invested in at positive and negative returns, but the probability of being invested in at the zero return remains unchanged, and the expected return on being invested in decreases.

Another necessary condition of equilibrium (x^*, m_l^*) is that the manager prefers x^* to any $x \in (x^*, 1]$

$$\begin{aligned} (x^*A + \beta) & \left[p(0) \int_{\underline{a}}^{\hat{a}(x^*)} 1dF(a) + p(\bar{R}) \int_{\underline{a}}^{\hat{a}(x^*)} 1dF(a) + p(\underline{R}) \int_{\underline{a}}^{\hat{a}(x^*)} 1dF(a) \right] \\ & \geq (xA + \beta) \left[p(0) \int_{\underline{a}}^{\hat{a}(x^*)} 1dF(a) + p(\bar{R}) \int_{\underline{a}}^{\hat{a}(x)} 1dF(a) + p(\underline{R}) \int_{\underline{a}}^{\hat{a}(x)} 1dF(a) \right]. \end{aligned}$$

Again, here we use the fact that the value of $m(\hat{a}, 0)$ does not influence the value of the integral.

C Institutional Background

This appendix provides examples of real-world intermediaries in delegated asset management, with a particular focus on personalized advisories and/or recommendation signals.

As new entrants to the asset management business, fintech platforms are the ones that most closely correspond to our baseline settings. They prioritize maximizing and protecting users' investments. They have been investing considerable effort in improving their technology and designs to provide better personalized services. For example, **Yieldstreet**, founded in New York in 2015, has grown up to a large business scale with more than 450,000 users and \$3.9 billion invested value up to September 2024.²⁰ It connects retail investors with alternative investments managed by various fund managers, offering personalized recommendations based on user profiles and investment goals. Figure C1 illustrates the three main steps involved in the platform. First, the platform recommends investment opportunities, highlighting its advantages in offering a wide range of options (including real estate, art and legal finance) and providing easy access. Second,

²⁰According to <https://www.yieldstreet.com/about/> Date of visit: Sep 09, 2024.

it helps users to *invest* according to their risk tolerance and the projects' past performance. Third, the platform tracks the *earnings* and especially visualize asset allocations.

Commercial banks have a long-standing tradition of providing asset management services. Some businesses have been split into specialized companies or platforms. In this era of digital finance, easy access, low costs for online/in-app usage and personalized services are commonly highlighted. For example, **Merrill Guided Investing** (under Bank of America), as shown in Figure C2 provides both automated investing and guided advisory services. The platform integrates human advisors with digital tools, allowing investors to choose from curated portfolios managed by fund managers. Merrill Lynch fund managers and other partner funds can promote their strategies on the platform, and users receive personalized recommendations based on their goals.

Similarly, **Wells Fargo Intuitive Investor** combines robo-advisory services with access to financial advisors and a marketplace for managed funds. Fund managers can promote their funds on the platform, and users receive investment recommendations based on their profiles and risk tolerance, as Figure C3 shows.

Since 2012, China has permitted platforms to distribute mutual funds. Technology companies that are independent of fund families, banks and brokers are allowed to issue mutual funds via fintech platforms. One of the largest platforms, Ant Financial, is a typical example of a platform that assesses investors' risk tolerance and investment goals and then recommends suitable funds.

Ant's ecosystem comprises five major components: online consumption, mobile payment, investment, consumer credit, and healthcare insurance. For a detailed introduction, see for example **Hong et al. (2020)**. This all-in-one ecosystem enables it to conduct analysis of investor preferences, as shown in Figure C4. The app interface features mutual fund recommendation pages, as illustrated in Figure C5.

Furthermore, Ant Financial has partnered with Vanguard Group to develop a fund investment advisory service called "BangNiTou". The system evaluates an individual's daily consumption, financial habits, and other data to create a personalized investment strategy based on their risk assessment results. This includes determining investment objectives, asset allocation, and expected returns. After the risk assessment and setting of investment goals, BangNiTou works by recommending a portfolio selected from 6,000 mutual funds, see Figure C6. In 2021, assets under BangNiTou sits at ¥6.9 billion (about \$1 billion) (Bloomberg, 2021). BangNiTou adopts a "buyer's agent" model, customizing financial planning based on an investor's risk assessment and a curated pool of mutual funds. The platform charges a service fee of approximately 0.0014% of total daily assets (0.5% annually). Fees related to fund transactions are charged according to the pricing rules of the respective fund products.

MITIGATING MORAL HAZARD THROUGH ALGORITHMS

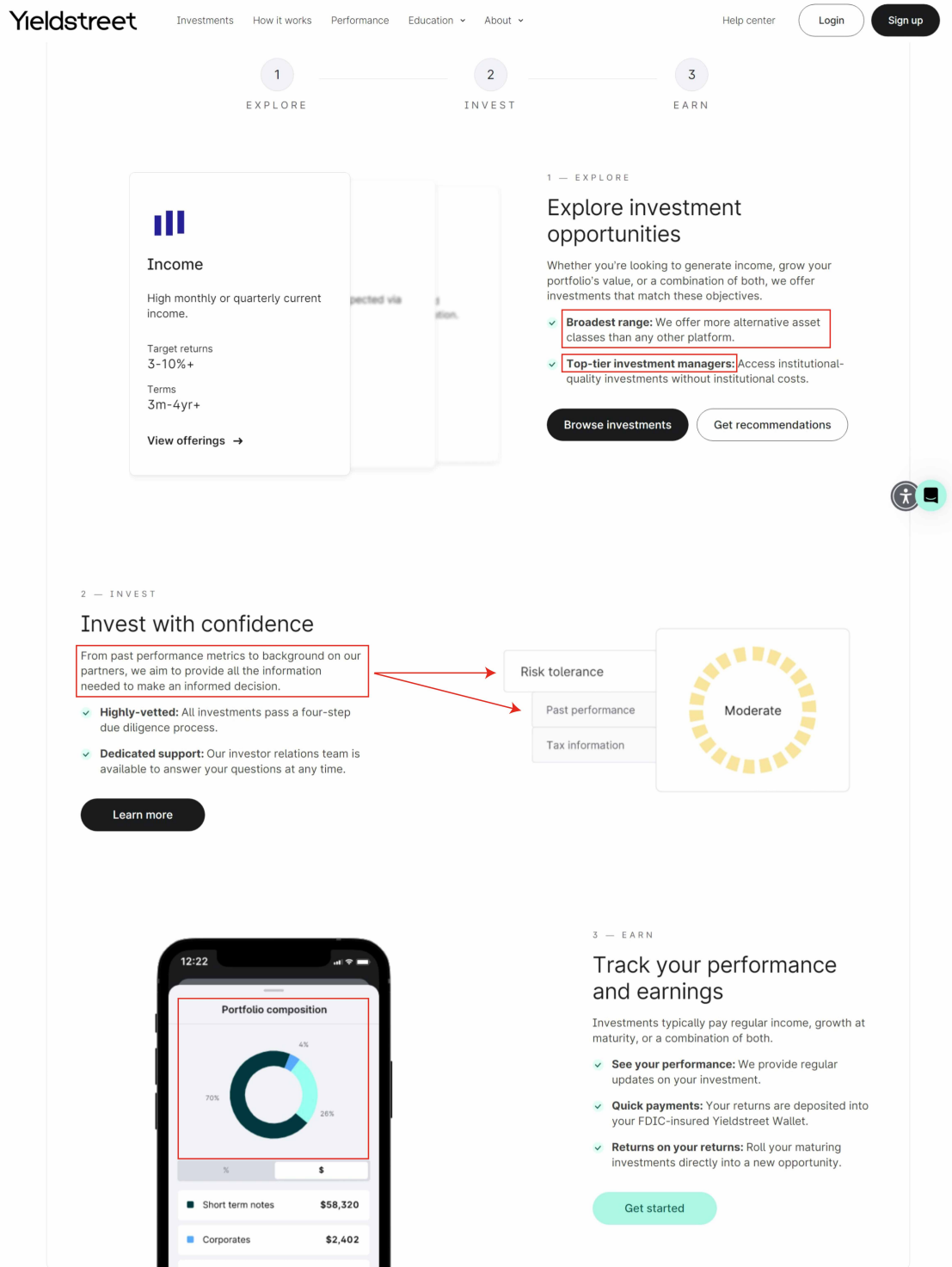


Figure C1: Delegated Investment service of Yieldstreet

Source: <https://www.yieldstreet.com/how-it-works/>. Date of visit: Sep 09, 2024.

MITIGATING MORAL HAZARD THROUGH ALGORITHMS

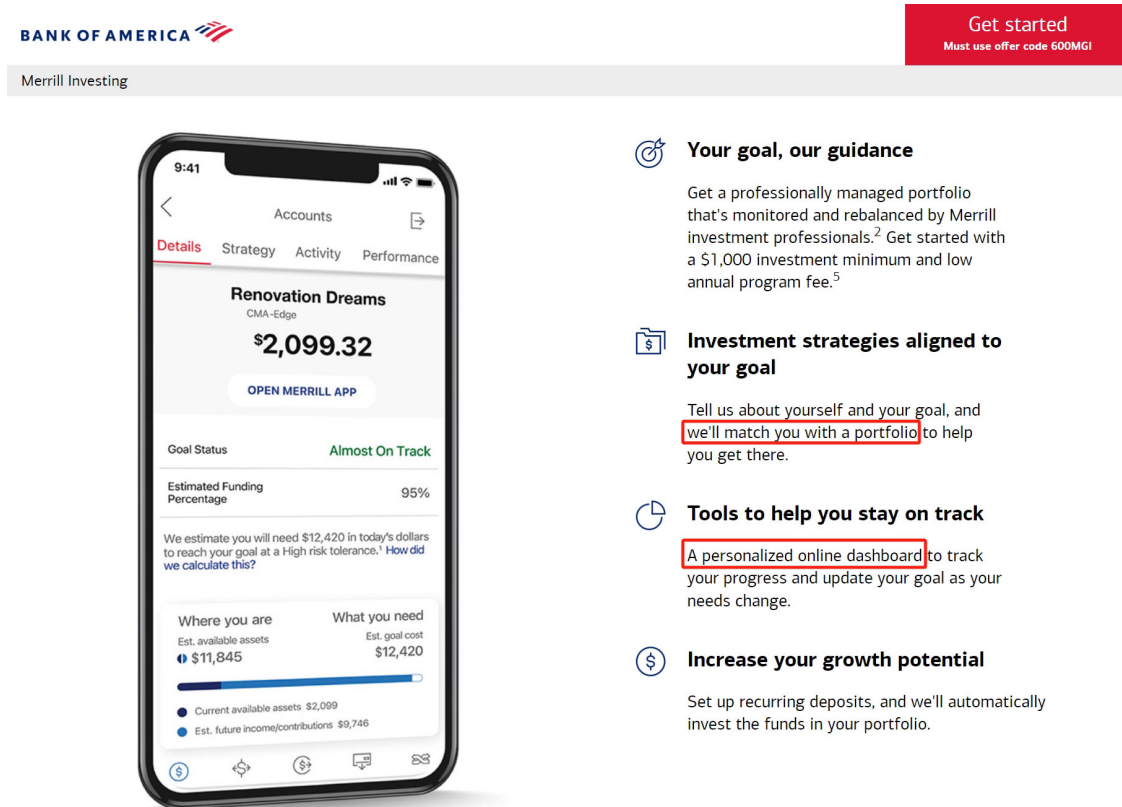


Figure C2: Personalized investment matching service of Merrill Guided Investing

Source: <https://www.merrilledge.com/offers/retirement-mgi>. Date of visit: Sep 09, 2024.



With Intuitive Investor® you get:

- A low cost, professionally designed portfolio
- Automated investing technology
- Access to financial advisors
- Goal tracking with LifeSync® to set and track progress toward your financial objectives

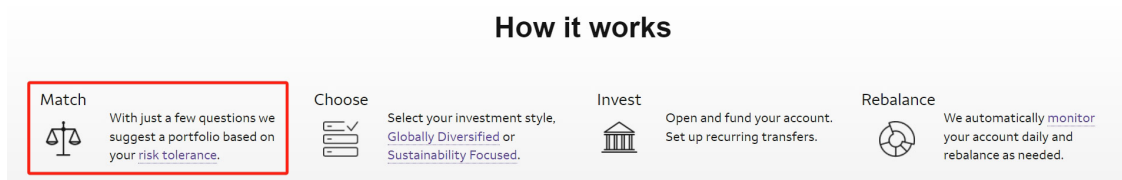


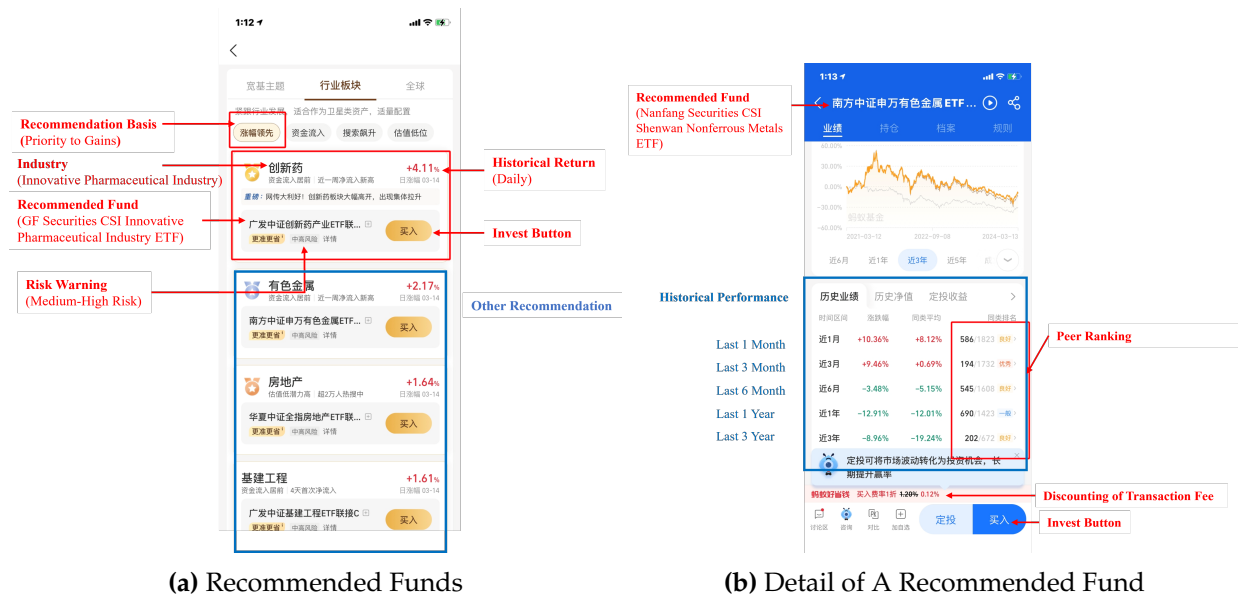
Figure C3: Intuitive Investor's Personalized investment based on risk tolerance

Source: <https://www.wellsfargoadvisors.com/services/intuitive-investor.htm>. Date of visit: Sep 09, 2024.

MITIGATING MORAL HAZARD THROUGH ALGORITHMS



Figure C4: Analysis of Risk Tolerance and Investment Object in Ant Financial

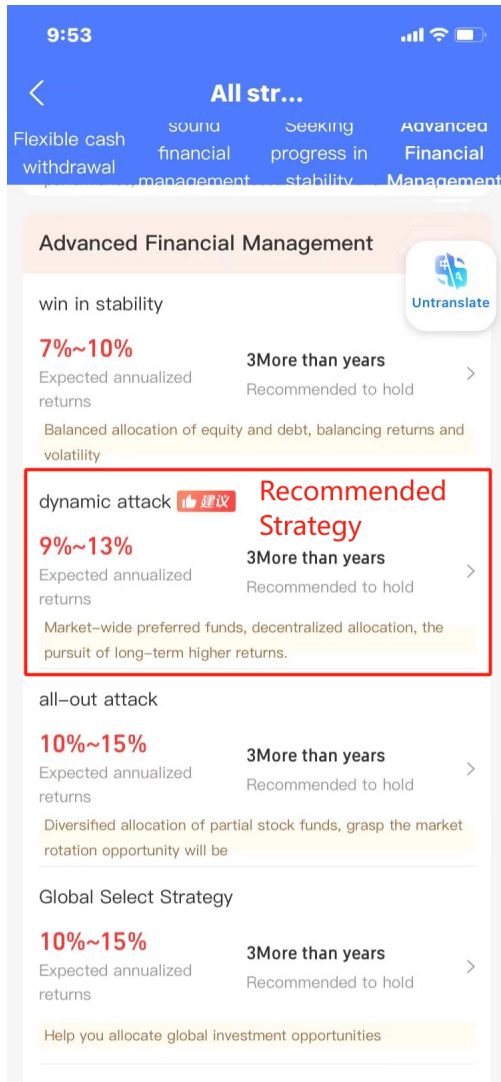


(a) Recommended Funds

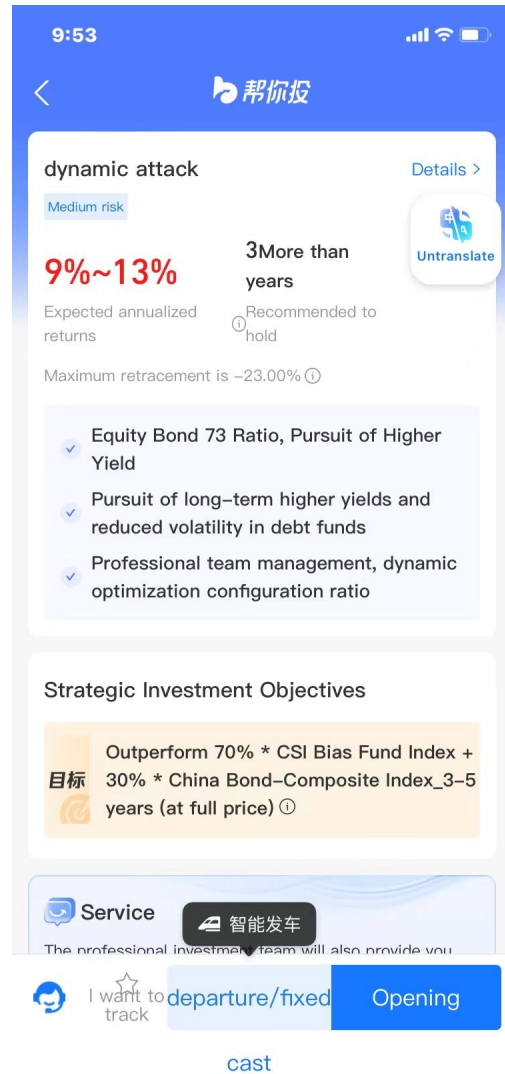
(b) Detail of A Recommended Fund

Figure C5: Recommendation service of Ant Group Co. Ltd.

MITIGATING MORAL HAZARD THROUGH ALGORITHMS



(a) Recommended Strategy



(b) Detail of the Strategy

Figure C6: Robo-advisory service of BangNiTou