Mitigating Moral Hazard in Delegated Investment through Recommendation Algorithms *

Kai Feng† Zhiheng He † Wenshi Wei †

December 18, 2024

It is growing rapidly that platforms serve as intermediaries in delegated asset management, especially featuring algorithms delivering personalized recommendations to a large user base. We develop a model to study the platform's investor-optimal algorithm design when investors contract with portfolio managers according to recommendation status and may be unaware of their risk aversion. We show algorithms can mitigate the manager's moral hazard in over risk-taking, which is inevitable under linear limited-liability contracts. The algorithm serves as an information gatekeeper with commitment power from leveraging the user base. In contrast with consumption platforms, the algorithm observes noisy historical performance, reduces recommendations under ambiguous signals, and potentially compensates for clear signals, generating information rent paid by investors. We provide an approach to solve general algorithm design problems and provide several extended discussions, including inequality concerns, joint design with contracts, comparisons with other delegated investment scenarios, and the co-existence with fund rating systems.

Keywords: algorithm design, moral hazard, delegated asset management, FinTech platform, risk chasing

*We are grateful to Bo Hu, Zanhui Liu, Dan Luo, Tat-How Teh, Yifan Tao, Longtian Zhang, and participants at conferences and workshops at the 2024 China Information Economics Society Annual Meeting, 2024 China Tech-Fin Research Conference, 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

Increasingly, we are coming to understand that financial technologies, which have been widely used in asset management, can remarkably reshape market participants and interactions (e.g., Cookson et al., 2021; Capponi et al., 2022; Hong et al., 2024). Be it played by banks or fintech platforms, institutions newly feature an intermediary that connects various investors and portfolio managers (e.g. mutual funds, private equity funds, insurance funds), aggregating a great market scale.[1] This expands financial inclusion by providing easy and cheap access to numerous investment advisories and opportunities. However, the concentration of investors implies a high potential of visibility on the platform. Limited-liability managers are then incentivized to increase risk-taking for standout performance, enhancing the exposure and sales (Hong et al., 2024). Their combination can be dangerous: numerous non-professional investors are overexposed to market risk, generating systemic risks to the market.

This paper analyzes why and how recommendation algorithm designs can address the above challenge. As platforms benefit from a large user base, they have the incentive to restrict the fund managers' risk-chasing and protect user welfare. There are three difficulties. First, non-professional investors typically have limited awareness of their heterogeneous levels of risk aversion (Capponi et al., 2022). This requires personalized intervention with private signals. Second, the platform cannot change the contracts between investors and managers, thus fails to directly mitigate the agency problem. Third, managers are always able to hide their allocation details.[2] Therefore, over risk-taking is always represented as a moral hazard.

We build a tractable model to show that in this context, all three difficulties can be simultaneously addressed by a recommendation algorithm. In the baseline, we consider a two-period economy where a continuum of non-professional investors contract with a fund manager through a platform. Investors have heterogeneous risk aversion and do not know their types. The risk-neutral manager designs a portfolio and keeps the fraction of risky-asset allocation as private information. The platform takes advantage of data, thus observing investors' risk aversion levels. The algorithm is a predetermined publicly known process that delivers personalized recommendation signals according to individ-

[1] These platforms typically allow access to both self-directed investing in funds and advisory services, such as Merrill Guided Investing owned by Bank of America, Citibank's Citi Personal Wealth Management, and Wells Fargo's Intuitive Investor etc. Emerging fintech platforms including Yieldstreet, iCapital Network, etc. The phenomenon is worldwide: 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 universe of mutual funds. See Appendix B for institutional background.

[2] We focus on active investment, where disclosure in practice is always delayed and not fully transparent. Also, as discussed later, the noisy signals from historical performance prevent the platform from correctly inferring the allocation.

ual risk aversion and the fund's historical performance. After observing the period-one portfolio return, the algorithm executes automatically. Investors update their beliefs about the risk aversion levels conditional on the recommendation status, and decide whether to invest in the fund, i.e., contract with the manager, to maximize the expected period-two payoff accounting for the delegation fee. Correspondingly, the manager designs the fund to maximize the expected total delegation earnings.

In practice, the delegation fee is commonly written in a simple contract, including a fixed management fee and a performance fee with limit liability. Then it is easy to expect that without an algorithm, the manager would optimally fully allocate in risky assets. However, with an algorithm, the manager knows that the realized return offers a noisy signal about the allocation choice, and the algorithm might punish the suspected over-risky performance by reducing recommendation probabilities. As a result, the manager optimally chooses a lower risk allocation under the algorithm. Specifically, we show that the platform is able to force the manager's equilibrium risk allocation to any targeted level, and further choose the equilibrium that optimizes the aggregate expected investor payoff. Proverbially speaking, the platform effectively controls the market participants and interactions by leveraging "the algorithm's hand."

In the above mechanism, a key question is: how can algorithms change the risk allocation of the manager? First, we emphasize that the critical role of algorithms is not limited to the transmission of information. Algorithms also provide commitment power. Through predetermined recommendation plan for all possible portfolio performance, algorithms provide punishment over excessive risk taking. This breaks the monotonicity of the manager's expected payoff with respect to the risk level of their fund.

Performance uncertainty has a subtle effect on algorithm implementation, as different realized returns can have various degrees of informativeness. To illustrate our key intuition, consider an over-risky portfolio with an overlapping return range with the target portfolio. In this case, algorithms cannot distinguish the manager's risk choice with full confidence when they observe a realized return that falls within the overlap. To counter such uncertainty, algorithms tend to be conservative in their recommendations on such overlap. Algorithms compensate otherwise, i.e., increase the probability of recommendation when observing informative signals from the non-overlapping return range of the target portfolio. Put differently, the manager has the incentive to hide behind noisy signals, while the algorithm rewards informative disclosure. Due to this trade-off, some investors receive recommendations that are not the first best, update their beliefs incorrectly, and suffer a welfare loss. Therefore, the compensation is effectively an information rent paid by investors. We further discuss its inference from the perspective of inequality: investors with high risk aversion levels are still first affected. They are sacrificed as the small for the greater good of the investors as a whole.[3]

We then analyze the relationship between contracts and algorithms. Whereas the contract designs incentive mechanisms between parties under a typically asymmetric information structure, the design of a recommendation algorithm reshapes the structure of information and participating parties. Here, recommended investors update their beliefs about risk aversion levels, and those not recommended may even not enter the market.[4] The contract affects the difficulty and effectiveness of implementing a recommendation algorithm. In particular, a higher fixed management fee aligns with the algorithm's goal. Under a sufficiently high management fee, the algorithm is able to optimize the expected aggregate investor payoff with zero information rent. Note that the management fee is also paid by the investors. As a result, we show that the aggregate effect generates an inverted U-shaped expected investor payoff with respect to the management fee. This sheds light on a jointly optimal design of algorithms and contracts.

Methodologically, we develop an approach for solving general algorithm design problems under continuously distributed returns, which is challenging despite the relatively well-understood intuitions. As illustrated, a high return is not always favored, but is more likely a sign of over risk taking. This forces us away from assuming monotonicity of the objective algorithm, thus invalidates a set of common approaches in mechanism design. In addition, the optimization works on function spaces, and few assumptions are made on the return distribution, which both add mathematical complexity.

Our approach consists of three main steps. First, in light of Ichihashi and Smolin (2023), we show that the possible optimal algorithm belongs to a specific family of functions featuring a simple threshold. Next, we consider three function spaces, provide corresponding sufficient conditions, and prove the existence of solutions to the general algorithm design problem, respectively. Finally, we use variational methods to derive the necessary conditions for the optimal algorithm, which restricts the interior solution to a unique analytical form. By pinning down any applicable return distribution, one can obtain a single algorithm through the necessary conditions and validate its satisfaction of the manager's incentive constraint, ultimately obtaining the optimal algorithm. We raise an example with a normally distributed risk return. Interestingly, the resulting algorithm exhibits an elegant quadratic shape. This approach provides a mathematical foundation for future research on a broad category of algorithm design problems.

[3] The underlying reason is that investors hold the same fund and cannot reallocate risk between themselves. If we consider multiple funds, then the algorithm can be designed to recommend different funds to heterogeneous investors, partially mitigating inequality issues.

[4] We assume in the baseline that investors are initially unaware of the fund, since in practice, there are too many funds to notice. For example, there are 7,222 mutual funds in the US in 2023 (referred to Statista Research Department), and 10,742 in China in 2024 (referred to Asset Management Association of China). We relax this assumption in the extended discussion, where all the core implications remain.

We extend to other relevant scenarios of delegated investment, thereby re-showing the unique roles of the algorithm. In a primordial case without platforms and algorithms, non-professional investors know neither the historical performance nor the risk aversion levels. They directly meet the manager and make decisions based on their beliefs about their own risk aversion levels, say, the population average level. The moral hazard is obviously serious: the "blind" investors have no information to infer the manager's choice. Step further, we allow the investors' access to the historical performance, e.g., investment on a traditional fund distribution platform, which only provides information about the fund, but do not make personalized services. The manager may be self-motivated to reduce risk exposure, since investors have more or less beliefs about their risk aversion. However, the effect on moral hazard is minimal, and investors may obtain ex-post negative payoffs.

Whereas the former comparisons emphasize the crucial personalized information, we further come to investment experts (or investment with advisory) to see that the algorithm has functionality beyond processing information. Investors now observe both the historical return and their personal characteristics. They fairly never get themselves over-exposed to risk and always obtain non-negative expected payoffs. However, the historical signals are not always informative, and the manager still has incentive to hide. The population cannot generate punishment, so let alone the compensation–each investor always refuses to invest with ex-post inefficiency and agree to invest with efficiency, generating no commitment power. Compared to our baseline with algorithms, the equilibrium yields a fund with higher risk exposure and a narrower inclusion. This lack of coordination, in a sense, exhibits "a curse of shrewdness," and contrasts the crucial role of algorithms in providing commitment power.

We further discuss the relationship and interactions between recommendation algorithms and fund ratings (e.g., Morningstar). They are typically motivated from the opposite sides of the platform: fund ratings compare and classify multiple funds, whereas algorithms target heterogeneous investors. They interact in the context of information structure, providing investors public and private signals, respectively. We analyze the implementation of algorithm with the co-existence of fund ratings, and notice the counterintuitive implication: the expected aggregate investor payoff generated by the optimal algorithm is reduced. This is because additional information makes the algorithm less determinative in investors' decision-making, results in a limited commitment power, and drives the possible equilibrium allocations in favor of the manager. The mechanism design combining fund ratings and recommendation algorithms calls for further exploration.

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 investment and expand adoption by introducing various technologies, such as easy access to the centralized information flow of 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). We start from the perspective that platforms have become intermediaries (e.g., Stoughton et al., 2011; Cookson et al., 2021), connecting 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 documents that new technologies generate new impacts on participants' behavior: the integration of daily consumption and investment activities has increased investors' risk-taking behavior (Hong et al., 2020), and the centralized information flow 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 of/given 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), in particular, generating risk-taking incentives (Stoughton, 1993; Lee et al., 2019).[5] Li and Tiwari (2009) solves an option-type bonus fee with an appropriate benchmark to overcome the moral hazard in risk choices, whilst a practical challenge is that, as emphasized by D'Acunto and Rossi (2021), there is a preference for offering simpler contracts to minimize the risk of operational errors by non-professional households. Be it costly or not to implement a relatively complex contract, the realistic context of intermediary asset management suggests that platforms are unable to adjust the contracts between investors and delegated managers. We show that, by influencing the information structure, the automated and personalized recommendation algorithm can successfully address moral haz-

[5] Existing 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.

ard of risk allocations inherent in simple contracts. This insight of using technology to refine contracts resonates with Cong and He (2019), which discusses how blockchain can expand the range of variables that can be written into contracts, thus finding a niche in finance for the function of blockchain and smart contracts. From a wider point of view of finance theories, the optimal algorithm design explores a novel interaction between contract and information design, whereas the combination mainly lies 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 mechanisms to empower buyers in a transaction. In a bilateral trade, Roesler and Szentes (2017) suggests that buyers can influence sellers' pricing strategies by acquiring incomplete information. Ravid et al. (2022) further refines the strategic interaction outcomes when buyers lack commitment power. Ichihashi and Smolin (2023) shows the recommendation algorithm can protect the total consumer surplus from personalized price discrimination. Xu and Yang (2023) consider the potential drawback of algorithms prioritizing consumers when agency problems arise from both sides. In contrast to bilateral trade, we study the moral hazard problem in purchasing financial services, which remarkably features noisy signals and uncertain payoffs.[6] We detail the underlying mechanism of the algorithm in processing information with different qualities. Information with low qualities becomes crucial for balancing investor payoff and the manager's incentive. This echoes the empirical evidence of ambiguity aversion in mutual fund investment (Li et al., 2017), and speaks Szalay (2005)'s theory implication, excluding compromising choices increases incentives for information acquisition, from another tale: the algorithm incentives allocation disclosure by reducing unsure recommendations.

Regarding the role of information transmission, the recommendation algorithm adds to the large literature on information gatekeepers (Baye and Morgan, 2001). Many studies show the various motives and roles 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 takes advantage of 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 targets the elimination of moral hazard.

[6]There are researches studying commitment in mutual fund investment from other perspectives, e.g., Huang et al. (2020) focuses on shaping the market reputation in repeated games.

2 The Model

Consider a two-period economy where users (non-professional investors) contract with a fund manager through a platform. The higher risk aversion of investors and manager's limited liability induce moral hazard, driving investors potentially over-exposed to market risk. The platform targets a large user base and therefore attempts to protect investor welfare. Despite inability to modify contracts between investors and the manager, the platform engages in their matching process by designing fund recommendation algorithms.

2.1 Setup

Fund manager. The risk-neutral fund manager designs a portfolio consisting of a risky asset and a risk-free asset, generating returns $R_{pt} = R_f + x(R_t - R_f)$, $t \in \{1, 2\}$, where the risk-free return R_f is normalized to zero without loss of generality. The risky return R_t is independently and identically distributed in the two periods. Denote $R_t \sim G$, where G is with a strictly positive density g over its support $[\underline{R}, \overline{R}]$. $\mathbb{E}[R_t], Var[R_t] < \infty$. The portfolio allocation is determined confidentially, thus the share of risky asset $x \in [0, 1]$ is the manager's private information. To the same reason, the platform and investors can only observe the realized historical portfolio return r_{p1} other than the realized risky return r_1, even though the platform knows the distribution G.[7]

The manager sells the fund on the platform at $t = 1$ to maximize their expected terminal utility. In particular, the financial payoff comes from the limited-liability delegated asset management contract, $\phi(r) = \max\{\alpha r, 0\} + \beta$, where the first term refers to a performance fee proportional to terminal wealth, $\alpha \geqslant 0$, and $\beta \geqslant 0$ is a fixed management fee. The manager may also be incentivized by personal benefits, e.g., becoming an attention-grabbing star with high performance, which also imposes asymmetric effects of the fund's gains and losses on the manager's utility.[8] For simplicity, we assume a similar form to the performance fee. The manager's expected terminal utility reads

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

[7]We focus on active funds where managers design specific portfolios to outperform the market and attract investors. The risky asset in our model de facto corresponds to a weighted combination of investible assets determined by the manager. In reality, the disclosure of fund positions is limited and lagging, leaving little certainty about x. We allow the platform to have knowledge about G, representing its outside advantage in processing market information, while investors may only obtain partial information, e.g., $\mathbb{E}[R_t], Var[R_t]$.

[8]This is similar to the private benefits received by entrepreneurs when they succeed in financing (e.g., Szydlowski, 2021).

where q is the total sale. The asymmetric revenue structure leads to agency problems: the manager would prefer to fully allocate in risky assets if neglecting the impact on sales. This is inevitable under a limited liability contract. As highlighted by Palomino and Prat (2003), it prevents investors from selling returns to the manager in exchange for their expected value.

Investors. A unit continuum of investors have heterogeneous risk aversion. They are indexed by their type a, where an a-type investor has \$1 for investment, decides whether to invest in the fund at $t = 1$ based on the expected quadratic utility over the terminal return:[9]

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

where a higher a captures a higher degree of risk aversion.[10] The distribution F of type a over the population has a strictly positive density f over its support $[\underline{a}, \overline{a}]$, $0 < \underline{a} < \overline{a}$. The key building block is that the investors are non-professional: they do not know their type a, and are not capable to search for a proper portfolio on their own. Instead, they decide whether to invest in the fund only after receiving its recommendation signal from the platform.

Platform and algorithm. Different with traditional mutual fund sales and ratings, the platform can leverage data and technologies to assess any investor's risk aversion and the fund's historical performance, and have an algorithm send *personalized recommendation signals*. Specifically, an algorithm is a function $m : [\underline{a}, \overline{a}] \times [\overline{R}, \underline{R}] \to [0, 1]$. For any pair of a and r_{p1}, the algorithm recommends each a-type investor the fund with probability $m(a, r_{p1})$.[11] The algorithm is *publicly* known to investors and the manager. In particular, the manager knows that the portfolio design would affect recommendation probabilities through historical performance, and further the total sales q.

Timeline. Formally, the timeline is as follows:

[9]Since a group of investors with the same risk aversion and endowments are identical in every respect, an a-type wealthy investor can be represented as a group of a-type investors with \$1 for investment. This allows wealth heterogeneity to be included through the distribution of a, whereas the \$1 endowment essentially assumes the total wealth for investment to be one, and is therefore without loss of generality.

[10]The expected utility also implies that the utility of not investing in funds is zero and is unaffected by risk aversion.

[11]Our framework allows for extension to multiple fund manager scenarios and competition. Regardless of whether a fund manager designs a differentiated portfolio, the platform can determine their competition structure via the algorithm, e.g., by recommending to separated groups or embedding additional ranking rules. This is beyond our prime interest.

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 $\phi(R_{p2})$, respectively.

2.2 Platform's Optimization and Solution Concept

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

The platform's problem is to solve an *investor-optimal algorithm*, which attains greater investors' total expected payoff than any other recommendation algorithm. This simplifies the problem as well as comes from the following two considerations. Regarding regulation, robo-advisors are considered fiduciaries under the Investment Advisers Act of 1940, which obligates them to act in the client's best interest (Capponi et al., 2022). As for incentives, Xu and Yang (2023) emphasizes that platforms, driven by the goal of maximizing future revenue, become "consumer-minded" since their market success relies on past consumer satisfaction. In practice, for example, BangNiTou, China's robo-advisor service, collects fixed fees from investors, not managers, and is therefore naturally accountable to investors.[12] As for investor-optimal algorithm, the optimization problem generates an outcome (x, m), i.e.,[13]

$$\max_{\substack{m:[\underline{a},\overline{a}]\times[\underline{R},\overline{R}]\to[0,1],\\ x\in[0,1]}} \int_{\underline{R}}^{\overline{R}}\int_{\underline{R}}^{\overline{R}}\int_{\underline{a}}^{\overline{a}} \left[(xr_2 - \phi(xr_2)) - \frac{1}{2}a\,(xr_2 - \phi(xr_2))^2\right] m(a, xr_1)\mathrm{d}F(a)\mathrm{d}G(r_2)\mathrm{d}G(r_1) \tag{1}$$

subject to the following constraints

$$x \in \arg\max_{x'} \left\{ \int_{\underline{R}}^{\overline{R}}\int_{\underline{R}}^{\overline{R}}\int_{\underline{a}}^{\overline{a}} \phi(x'r_2)m(a, x'r_1)\mathrm{d}F(a)\mathrm{d}G(r_2)\mathrm{d}G(r_1) \right\}, \tag{2}$$

[12] BangNiTou, developed by Ant Financial and Vanguard Group, adopts a "buyer's agent" model, customizing financial planning based on the investor's risk assessment and a pool of mutual funds. The service fee for "BangNiTou" is calculated as approximately 0.0014% of total daily assets (equivalent to 0.5% annualized) and is charged quarterly. Fees related to fund transactions are charged according to the pricing rules of the respective fund products.

[13] Note that the risk-free rate is normalized to zero, then $r_{pt} = xr_t$, $t = 1, 2$.

$$\int_{\underline{R}}^{\overline{R}} \int_{\underline{R}}^{\overline{R}} \int_{\underline{a}}^{\overline{a}} \phi(xr_2) m(a, xr_1) \mathrm{d}F(a) \mathrm{d}G(r_2) \mathrm{d}G(r_1) \geqslant 0, \tag{3}$$

$$\int_{\underline{R}}^{\overline{R}} (xr_2 - \phi(xr_2)) - \frac{1}{2} \frac{\int_{\underline{a}}^{\overline{a}} am(a, xr_1) \mathrm{d}F(a)}{\int_{\underline{a}}^{\overline{a}} m(a, xr_1) \mathrm{d}F(a)} (xr_2 - \phi(xr_2))^2 \, \mathrm{d}G(r_2) \geqslant 0, \forall r_1 \in \mathrm{supp}(R_1). \tag{4}$$

Eq. (2) is the manager's incentive-compatible constraint, i.e., the equilibrium allocation x maximizes the manager's expected financial payoff with the corresponding recommendation algorithm. Eq. (3) is the manager's IR constraint. It is naturally satisfied, because the algorithm design merely alters recommended probabilities, incorporating the specified limited liability contract $\phi(\cdot)$. Eq. (4) is the investors' IR constraint. For investors who receive recommendations, they 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 recommended fund is given by

$$\int_{\underline{R}}^{\overline{R}} \int_{\underline{a}}^{\overline{a}} (xr_2 - \phi(xr_2)) - \frac{1}{2} a \, (xr_2 - \phi(xr_2))^2 \, \mathrm{d}F(a|\text{recommended}, xr_1) \mathrm{d}G(r_2)$$
$$= \int_{\underline{R}}^{\overline{R}} (xr_2 - \phi(xr_2)) - \frac{1}{2} \mathbb{E}[a|\text{recommended}, xr_1] \, (xr_2 - \phi(xr_2))^2 \, \mathrm{d}G(r_2),$$

which obtains the R.H.S. of (4). An investor satisfying the IR constraint would prefer to contract if being recommended and updating their beliefs about their risk aversion. Put differently, this constraint binds 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

$$\int_{\underline{R}}^{\overline{R}} \int_{\underline{a}}^{\overline{a}} \left[(xr_2 - \phi(xr_2)) - \frac{1}{2} a \, (xr_2 - \phi(xr_2))^2 \right] m(a, xr_1) \mathrm{d}F(a) \mathrm{d}G(r_2) \geqslant 0.$$

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:

$$\mu_+ := \int_0^{\overline{R}} r_2 g(r_2) \mathrm{d}r_2, \qquad A := (\alpha + \gamma)\mu_+,$$
$$k_1(x) := \int_{\underline{R}}^{\overline{R}} (xr_2 - \phi(xr_2)) \mathrm{d}G(r_2), \quad k_2(x) := \int_{\underline{R}}^{\overline{R}} (xr_2 - \phi(xr_2))^2 \mathrm{d}G(r_2).$$

We make the following reasonable assumptions in solving the model:

Assumption 1.

1. *The investor's utility* $u_I(r)$ *is increasing and concave:* $1 - ar > 0$ *for all* $a \in [\underline{a}, \overline{a}]$ *and all* $r \in [\underline{R}, \overline{R}]$.

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 of the portfolio falls below that of a risk-free asset. Under Assumption 1, we have $\forall a \in [\underline{a}, \overline{a}]$,

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

3 Preliminary Analysis

Before rigorously analyzing the existence and implications of the optimal algorithm, this section uncovers its necessary properties: the optimal algorithm comes from a family of functions that features a simple threshold and requires a penalty for abnormal returns.

3.1 Threshold Algorithm

We note a special set of algorithms in light of Ichihashi and Smolin (2023)'s seminal work. Precisely, denote an algorithm m as a *threshold algorithm* if there exists a *threshold function* $\hat{a} : [\underline{R}, \overline{R}] \to [\underline{a}, \overline{a}]$ such that $m(a, r_{p1}) = \mathbb{1}(a < \hat{a}(r_{p1}))$. That is, a threshold algorithm recommends the fund with probability 1 (0) if the risk aversion is below (over) the threshold determined by historical returns.

Lemma 1. *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 cases under* m*.*

According to Lemma 1, suppose there exists an optimal algorithm m^*, we can always find a corresponding threshold algorithm $\hat{m}^*$ that ensures both the investor's IR condition and the manager's IC and IR condition are satisfied, while also ensuring that the investor's expected utility does not decrease. It suggests that we can find the investor-optimal algorithm in a set of threshold algorithms. The logic is simple: the manager views contracted investors as identical, leaving room for the algorithm to "select" proper investors to satisfy the manager's incentive constraint. In particular, for any given risky portfolio, investors with lower risk aversion always have higher expected utilities. Therefore, any selected (recommended) investor should have lower risk aversion than any unrecommended one, otherwise the total welfare can be increased by exchanging their recommendation states. As a result, Lemma 1 suggests the optimal algorithm (if exists) essentially determines the recommendation amount, and delivers to investors in order of their risk aversion.

Then we can 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 $[\underline{a}, \overline{a}]$, ranging from 0 to 1. Let $q := \int_{\underline{a}}^{\hat{a}} 1 \mathrm{d}F(a) = F(\hat{a})$, where $\hat{a}$ is affected by the realized historical return $r_{p1} = xr_1$.[14] With $q(xr_1) : [\underline{R}, \overline{R}] \to [0,1]$ 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

$$\max_{\substack{q:[\underline{R},\overline{R}]\to[0,1],\\ x\in[0,1]}} \int_{\underline{R}}^{\overline{R}} k_1(x)q(xr_1) - \frac{1}{2}\left(\int_{\underline{a}}^{F^{-1}(q(xr_1))} a\mathrm{d}F(a)\right) k_2(x)\mathrm{d}G(r_1) \tag{5}$$

subject to

$$x \in \arg\max_{x'} \left\{ (Ax' + \beta) \int_{\underline{R}}^{\overline{R}} q(x'r_1)\mathrm{d}G(r_1) \right\} \text{ and} \tag{6}$$

$$k_1(x)q(xr_1) - \frac{1}{2}\int_{\underline{a}}^{F^{-1}(q(xr_1))} a\mathrm{d}F(a)k_2(x) \geqslant 0, \forall r_1 \in \mathrm{supp}(R_1). \tag{7}$$

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

We further equivalently represent the IR constraint in a simpler form. Multiply the both sides of (7) with $q(xr_1)$ and focus on the left 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 with $q(xr_1)$. Also note (7) takes the equal sign when $q(xr_1) = 0$. Then we can define $\overline{q}(x)$ as

$$\overline{q}(x) := \sup \left\{ q \in [0,1] \middle| k_1(x)q - \frac{1}{2}\int_{\underline{a}}^{F^{-1}(q)} a\mathrm{d}F(a)k_2(x) \geqslant 0 \right\},$$

and the IR constraint is equivalent to

$$q(xr_1) \leqslant \overline{q}(x), \quad \forall r_1 \in [\underline{R}, \overline{R}].$$

[14] Recall that R_f is normalized to zero, and r_1 is not separately observed.

Briefly, investors would not buy if the platform over-delivers signals.

3.2 High Historical Returns Are Not Always Favored

The primary principal-agent problem here is the manager's tendency to overinvest in risky assets driven by their expected utility, which increases monotonically with x. Then the algorithm is designed to change such monotonicity and concavity 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.

Proposition 1. *If $q(\cdot)$ characterized by an algorithm is weakly increasing with $r \in [\underline{R}, \overline{R}]$, then it induces $x^* = 1$ in equilibrium.*

Intuitively, if $q(\cdot)$ weakly increases with r, then the algorithm provides the same incentives as contracts $\phi(\cdot)$ to the manager. As a result, the manager fully invests in risky assets to maximize the expected return–the algorithm fails to bind the manager's over-exposure to risk, although it could still prevent investors from entering the market who have negative expected utility under $x = 1$.

An important corollary is: the platform should reduce recommendations when historical returns are abnormally high, rather than viewing them as outstanding projects. 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). Algorithms contribute to mitigating this principal-agent problem. Importantly, compared to ranking systems, the algorithm allows for the integration of bilateral data from fund managers and investors, with personalized recommendations conveying private signals. In other words, the ranking system serves as a comparison metric of funds, while the algorithm focuses on evaluating a fund's suitability for a particular investor, and therefore enables a more direct penalty for risk-chasing. 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 (almost) impossible under an equilibrium allocation. Therefore, we can further characterize a feasible form of the optimal algorithm, detailed in the proposition below.

Proposition 2. *For any equilibrium (x^*, q^*), $\hat{q}$ constitutes a cutoff algorithm in an equilibrium $(x^*, \hat{q})$, generating the same expected payoffs of the investors and manager as (x^*, q^*), where*

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

Intuitively, $\hat{q}$ implements a greater penalty than q^* once the realized return exceeds $\text{supp}(x^* R_t)$, as these abnormalities imply the manager's deviation from the incentive compatibility. On the other hand, $\hat{q}$ imposes no additional penalty in equilibrium x^*. This ensures that $(x^*, \hat{q})$ still satisfies the IC constraint without changing the equilibrium investor utility at x^*. 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 focus on the existence of the equilibrium (x, q) and analyze the implications of the optimal algorithm featuring the form $\hat{q}$ without loss of generality.

4 Analysis under Discrete Return

This section develops the main intuitions with a discrete risk return. Section 5 considers the general case with continuous risk return. Here, we assume a follows a uniform distribution, and $\text{supp}(R_t) = \{\underline{R}, 0, \overline{R}\}$, reflecting the period-2 states "down", "flat", and "up", with probabilities $p(\underline{R})$, $p(0)$, and $p(\overline{R})$, respectively.

Given the discrete distribution, the optimization problem (5) can be transformed into solving the allocation x^* and three points $\hat{q}(x^* R_t)$ according to Proposition 2. As a result, the existence of the investor-optimal algorithm can be guaranteed by convex optimization on a compact set. In addition to the ease of solution, the discrete distribution also helps us to elaborate on the insights of the extension in Section 6.

4.1 Moral Hazard Mitigation and Information Rent

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 $r_1 \neq 0$, as it knows $r_1 \in \{\underline{R}, \overline{R}\}$ and also observes $r_{p1} = x r_1$. 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. Therefore, the key to mitigate moral hazard is to introduce a penalty under the uninformative flat case: the platform is expected to conservatively narrow the recommendation delivery to protect investor welfare. Then how to generate enough expected sales to achieve an ex-ante incentive compatibility of the manager? As compensation, the platform slightly expands its recommendations when it has information on x, even including investors with insufficient risk tolerance. This brings in welfare losses more or less, which is essentially an information rent paid to the manager.

To show the trade-off between cases of an algorithm, one could imagine the "ideal" (social planner's) ex-post recommendation, where all the investors with non-negative expected utilities under allocation x are recommended. It can also be understood as the first-best recommendation strategy assuming away the manager's incentive compatibility condition, i.e., $q_{FB}(x) = k_1(x)/(k_2(x)/2)$. Proposition 3 rationalizes the above intuitions.

Assumption 2. *The manager is disinclined to invest all capital in risk-free assets:* $(A+\beta)p(0) > \beta$.

Proposition 3. *Under Assumption 2, consider any equilibrium* $(x^*, \hat{q}^*)$ *where* $\hat{q}^*$ *characterizes a cutoff algorithm defined in Proposition 2 and it satisfies* $\max_r\{\hat{q}^*(r)\} > 0$. *Denote*

$$\underline{x} := p(0) - (1-p(0))\beta/A.$$

If $x^* \geqslant \underline{x}$, $\hat{q}^*(x^*\underline{R}) = \hat{q}^*(0) = \hat{q}^*(x^*\overline{R}) = q_{FB}(x^*)$. *Otherwise,* $\min\{\hat{q}^*(x^*\underline{R}), \hat{q}^*(x^*\overline{R})\} \geqslant q_{FB}(x^*) \geqslant \hat{q}^*(0)$. *When* $q_{FB}(x^*) \in (0,1)$, *the inequalities hold strictly.*

Proposition 3 highlights a crucial threshold of the exposure to risk $\underline{x}$. When an algorithm targets an allocation x beyond this threshold, it mitigates moral hazard without paying any information rent, acting as a de facto social planner. Economically, $x \geqslant \underline{x}$ is equivalently represented as

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

where the left side is the manager's expected payoff of good behavior from one deal, i.e., allocate x risky assets that align with the platform's target. The right side the opportunity cost of being good, i.e., the maximum expected payoff of deviating towards over risk-taking: if so, the manager would alternatively choose a full allocation in risky assets, and have a probability $p(0)$ to survive and get recommended.[15] Thus $x \geqslant \underline{x}$ guarantees the incentive compatibility without additional rent.

Otherwise, the platform needs extra efforts to push the incentive constraint: by penalizing the flat case and subsidizing other cases in terms of recommendation counts. In general, when the algorithm aims to reach a lower x, the manager is more inclined to incentive incompatibility and requires more compensation. The more information rent to pay further diverges the ex-post cases with different realized returns.

In sum, despite potential rents, the algorithm successfully enables the platform to "choose" a proper equilibrium allocation x, thus protecting the aggregate investor welfare. It also speaks the importance of an algorithm's commitment power, i.e., preemptively set recommendation probabilities based on varying realized performance with ex-post errors.

[15]Note that according to the definition of $\hat{q}(r;x)$, for any allocation $x' > x$, the fund would not be recommended if $r_1 \in \{\underline{R}, \overline{R}\}$. $p(0)(Ax'+\beta)$ increases in x', therefore the manager would choose $x' = 1$ once deviate.

Therefore, moral hazard mitigation relies on transparent algorithms and commitment to enforcement.

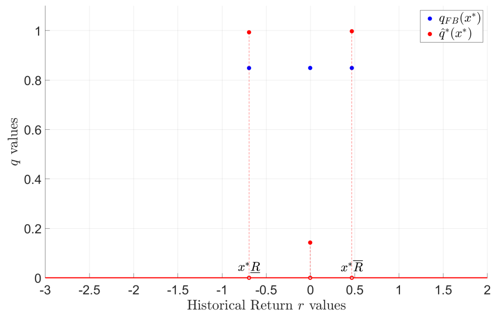
Figure 1 visualizes the optimal algorithm $\hat{q}^*$ and optimal allocation x^*. Panel (a) shows the optimal algorithm $\hat{q}^*$ that achieves an investor-optimal equilibrium x^*, as illustrated in Panel (b). The primary observation is that the algorithm punishes the manager by delivering no recommendation at all the realized returns that are impossible under x^*. The algorithm also punishes the flat-price case, but leaving a non-zero recommendation, as it still has a probability to be obtained from x^*. In addition, the gaps compared to the first-best demonstrate how the algorithm pays the information rent to the fund manager, thereby ensuring incentive compatibility.

Panel (b) illustrates the expected aggregate investor utility, which draws an inverted U-shaped curve, while for the no-algorithm case, the only equilibrium is $x = 1$ and investors receive non-positive expect utility. Put differently, the U-shape curve already speaks the power of algorithm–it enables the platform to choose any allocation as an aimed equilibrium, and optimizing investors' utility despite potential rent to pay. In other words, the intervention of algorithms removes the manager's full control over the allocation. Furthermore, the U shape owes to that the algorithm makes recommendations by Lemma 1 in an order according to risk aversion, whilst for each recommended investor, the expected utility goes up and then down as risk exposure increases and exceeds their risk tolerance. Especially, the downward interval implies the manager's increasing risk-chasing gradually generates more harm to investors. In addition, Panel (b) also shows the underperformance of the algorithm relative to the ideal case when $x < \underline{x}$, due to the required information rent and corresponding ex-post errors.

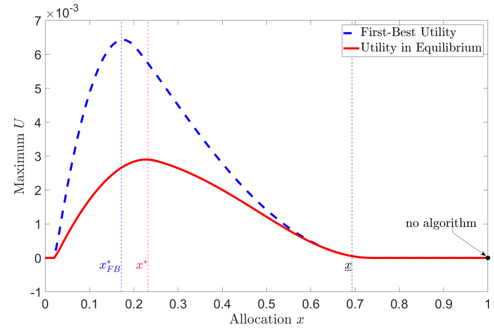
4.2 Algorithm Implementation and Inequality

Who actually pays the information rent? In this economy, the platform transfers the rent to investors. As Figure 1 Panel (c) shows, the ideal ex-ante recommendation q_{FB} is naturally determined by x and irrelevant to historical performance. However, the algorithm's recommendation varies in different period-1 states. When the fund exhibits a flat state, the algorithm makes an insufficient recommendation scale relative to q_{FB}. That is, some investors who were objectively eligible to invest are not recommended, resulting in foregone welfare gains, as depicted by Area B. In contrast, when the risky asset generates a non-flat state, the algorithm recommends an additional population with negative expected payoffs, as depicted by Area A.[16]

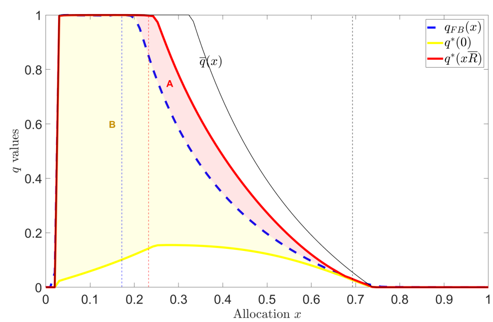
[16] Note that these investors voluntarily follow the recommendation guaranteed by the IR constraint: they obtain non-negative expected payoffs based on their subjective posterior risk aversion conditional on being recommended.



(a) $\hat{q}^*$ and q_{FB} for optimal allocation x^*



(b) $U(x, \hat{q}^*)$ for different x



(c) $\hat{q}^*$ for different x

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^*. 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 $R_{p1} \in \{x^*\underline{R}, x^*\overline{R}\}$. Outside $\text{supp}(x^* R_1)$, the algorithm sets the recommendation probability to 0 as Proposition 2. Panel (b) shows expected utility of investors under optimal incentive-compatible and first-best algorithms, given different x. Panel (c) shows the optimal incentive-compatible and first-best algorithms given different x. Area A and B represent the over-recommended and under-recommended investor populations respectively which results in the welfare gap in Panel (b). 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.

That is, in any case, there are always individuals who become "sacrificed" to the algorithm. In particular, according to the definition of the threshold algorithm, detailed in Lemma 1, the sacrificed investors lie between the ideal and true thresholds, especially the lower risk-tolerant ones than the unaffected and recommended ones. A low risk tolerance may correspond to tight budget constraints or less expertise. Therefore, inequality could deepen when the platform counters the fund manager with an algorithm.

Should it be that the algorithm inevitably has potentials to victimize the marginal users, even if they are supposedly optimal? We emphasize this to be sure, as long as the algorithm intends to eliminate moral hazard.[17] First, the platform's bargaining power comes from its user base as a whole, whereas any individual is atomic that has minimal influence to the manager's allocation choices. Therefore, its leverage in negotiations should stem from the user base. Second, the platform mitigates moral hazard by rewarding the manager's disclosure on x (although passive). The exact approach is to expand the recommendation scale, which may result in collective over-investment (relative to the collective risk aversion). Third, the recommendation does not redistribute risk exposure, therefore the collective over-investment manifests some individuals' over-investment. The lower aimed x requires more costly implementation and ultimately exceeds its contribution to welfare protection. As a result, one can again expect x^* to be larger than the optimal risk exposure under the same recommendation probability under equilibrium.

Proposition 4. *For any case where the equilibrium allocation exceeds zero under the optimal algorithm, i.e.,* $x^* > 0$*,*

$$\sum_{r \in supp\{R_t\}} p(r)\hat{q}^*(x^*r) \leqslant q_{FB}(x^*) = \min\{k_1(x^*)/(k_2(x^*)/2), 1\} \text{ and}$$

$$x^* \geqslant \sup\left\{\arg\max_{x'}\left\{k_1(x') \sum_{r \in supp\{R_t\}} p(r)\hat{q}^*(x^*r) - \frac{1}{2}\sum_{r \in supp\{R_t\}}\left(\int_{\underline{a}}^{F^{-1}(\hat{q}^*(x^*r))} a \mathrm{d}F(a)\right) k_2(x')\right\}\right\}.$$

In particular, when $x^* \in (0, \underline{x})$*, the inequality signs hold strictly.*

Proposition 4 uncovers the deviations of expected participation and risky asset allocation relative to the social planner's solution without the incentive-compatible constraint. The reason is that algorithmic designs must account for information rent costs as outlined in Proposition 3. 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 lowered. To ensure the manager is incentive compatible, the algorithm concedes in risk exposure x^*.

4.3 Interaction with Contract

Proposition 3 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

[17] If the algorithm were not designed to influence the manager, it would be easy in practice to prevent any investor from a welfare loss, e.g., by adopting an extremely conservative recommendation and focusing on $x \geqslant \underline{x}$.

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 3 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.

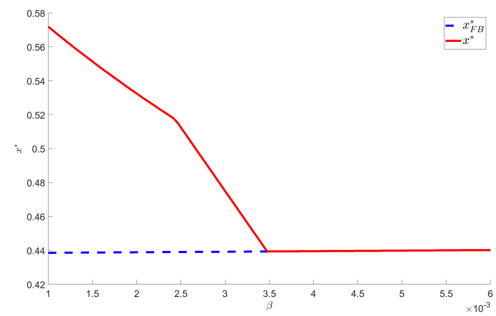
Figure 2 presents a comparative static analysis on β. As Panel (a) and (b) show, when the management fee is low, the equilibrium allocation under optimal algorithm, x^*, is higher than x^*_{FB}, and the expected recommendation scale is lower.As β increases, x^*_{FB} remains relatively stable, while x^* 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,[18] i.e., $x^* > \underline{x}$ as outlined in Proposition 3.

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

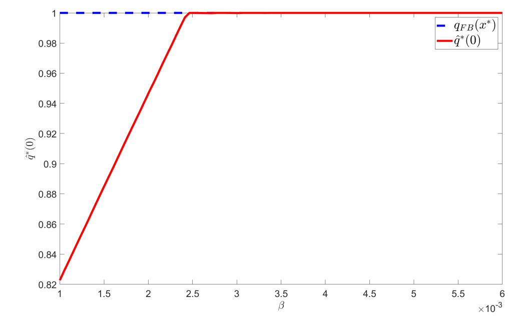
$$\mathbb{E}[u_I(x, \hat{q}^*(x))] = \underbrace{k_1(x)q_{FB}(x) - \frac{1}{2}k_2(x)\int_{\underline{a}}^{F^{-1}(q_{FB}(x))} a\mathrm{d}F(a)}_{\substack{\text{First-best utility,} \\ U(x, q_{FB})}} - \underbrace{\left| k_1(x)\Delta_q(x) - \frac{1}{2}k_2(x)\Delta_a(x) \right|}_{\substack{\text{Utility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

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 in 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 recom-

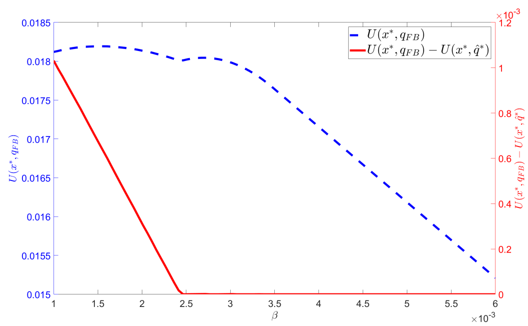
[18] The threshold of a sufficiently-high β is (about) 0.0035 under the parameter choice of Figure 2.



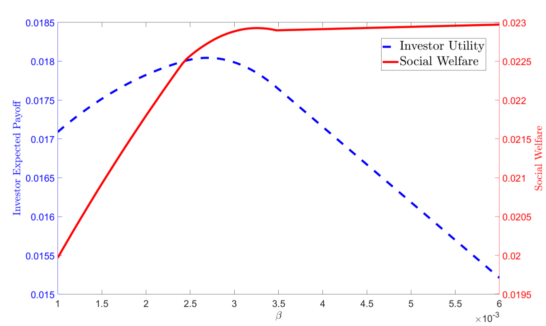
(a) x^*_{FB} and x^* for different β



(b) $q_{FB}(x^*)$ and $\hat{q}^*(0)$ for different β



(c) Decomposition of Utility



(d) Investor Utility and Social Welfare

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. By comparing with the First Best, as β rises, moral hazard diminishes and the algorithm $\hat{q}^*$ and allocation x^* move closer to the First Best (see Panels (a) and (b)). Considering that an increase in β directly leads to a loss in investor utility, the investor's utility initially increases with β but later declines (see Panels (c)). 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.

mended investors,

$$\Delta_q(x) = \underbrace{-p(0)(q_{FB}(x) - \hat{q}^*(0))}_{\text{Under-recommendation}} + \underbrace{\sum_{r \in \{\underline{R}, \overline{R}\}} p(r)(\hat{q}^*(xr) - q_{FB}(x))}_{\text{Over-recommendation}},$$

$$\Delta_a(x) = \underbrace{-p(0) \int_{F^{-1}(\hat{q}^*(0))}^{F^{-1}(q_{FB}(x))} a \mathrm{d}F(a)}_{\text{Under-recommendation}} + \underbrace{\sum_{r \in \{\underline{R}, \overline{R}\}} p(r) \int_{F^{-1}(q_{FB}(x))}^{F^{-1}(\hat{q}^*(xr))} a \mathrm{d}F(a)}_{\text{Over-recommendation}}.$$

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 (although 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.

5 Analysis under Continuous Distribution

In the previous section, uncertainty arises from the three possible future states. It 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 (albeit different) of yielding any specific historical returns. Therefore, (i) any historical return fails to fully infer the allocation, and (ii), even a zero return can be partly informative. Then the algorithm is expected to have strictly positive recommendation amounts on a continuous support, where at each point the recommendation decision accounts for the actions under all the possible allocations.

In this section, we characterize the implications when $\text{supp}(R_t) = [\underline{R}, \overline{R}]$. The predetermined algorithm infers the manager's choices based on realized historical returns with varying confidence, and enforces different recommendations accordingly. Then one can imagine that the previous implications still hold: the algorithm implements punishment, i.e., delivers conservative recommendation, when receiving less-informative and/or dangerous signals, and potentially pays an information rent. Ultimately, it 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 the optimal algorithm is paramountly not trivial because of the lack of the typical monotonicity. Mathematically, Proposition 1 implies that Helly's selection theorem, commonly used in the literature on mechanism or contract design, cannot be directly applied. For a rigorous definition of the existence problem, we constrain $q(\cdot)$ to the intersection of $L^\infty((\underline{R},\overline{R}))$ and function space $\mathcal{Q}$. Then the optimization problem (5) can be generalized as

$$\sup_{(x,q)\in D\cap[0,1]\times\mathcal{Q}} O(x,q),$$
$$D := \left\{(x,q) : q \in L^\infty\left(\left(\underline{R},\overline{R}\right)\right), 0 \leqslant q(\cdot) \leqslant \overline{q}(x) \text{ and } (x,q) \text{ satisfies (6)}\right\},$$
$$\overline{q}(x) := \sup\left\{q \in [0,1] \middle| k_1(x)q - \frac{1}{2}\int_{\underline{a}}^{F^{-1}(q)} a\mathrm{d}F(a)k_2(x) \geqslant 0\right\},$$

where $\overline{q}(x)$ represents the maximum possible fraction q under allocation x.[19] That is, $q(xr_1) \leqslant \overline{q}(x), \forall r_1 \in [\underline{R},\overline{R}]$. It can be proved that $\overline{q}(x)$ is continuous with respect to x.[20]

A reasonable choice of the function space $\mathcal{Q}$ ensures a certain sequential compactness of the feasible set and thus avoids difficulties that arise from not being able to constrain $q(\cdot)$ to be monotonic. For example, constraining $q(\cdot)$ to a Lipschitz function space with the same constant L, mathematically guarantees the continuity of the objective function and thus 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. *Assume that F is supported on $[\underline{a},\overline{a}]$ with continuous density function $f > 0$, where $-\infty < \underline{a} < \overline{a} < +\infty$. Assume also G has continuous density function $g > 0$ on $\left[\underline{R},\overline{R}\right]$. Suppose $\bar{q}(\cdot)$ is continuous. Let $L > 0$ be a constant. If one of the following three conditions applies,*

1. $\mathcal{Q} = \left\{q \in C_b\left(\left(\underline{R},\overline{R}\right)\right) : 0 \leqslant q(\cdot) \leqslant 1 \text{ and } \sup_{x\neq y} \frac{|q(x)-q(y)|}{|x-y|} \leqslant L\right\}$;
2. $\mathcal{Q} = \left\{q \in W^{1,p}\left(\left(\underline{R},\overline{R}\right)\right) : 1 < p < \infty, 0 \leqslant q(\cdot) \leqslant 1 \text{ and } \|Dq\|_p \leqslant L\right\}$;
3. $\mathcal{Q} = \left\{q \in BV\left(\left(\underline{R},\overline{R}\right)\right) : 0 \leqslant q(\cdot) \leqslant 1, q(0) = 0 \text{ and } \|Dq\|_{TV} \leqslant L\right\}$.[21]

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

[19] Note that the historical performance affects fraction of recommended investors $q(xr_1)$, whereas the manager determines the allocation x before the realization of R_1.

[20] The continuity is essential to demonstrate that a solution exists for the designer's optimization problem.

[21] Under *condition 3.* we always identify $q \in \mathcal{Q}$ with its version such that $q(0) = 0$. This is possible since the total variation defined as (8) is invariant under changes on sets of measure zero.

Remark 1. The notations and the basic concept of the proof are outlined as follows. Those not engaged in detailed mathematical analysis may choose to 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, and $\|Du\|_{TV;\Omega}$ denotes the *total variation* of the $L^1(\Omega)$ function $u : \Omega \to \mathbb{R}$, i.e.,

$$\|Du\|_{TV;\Omega} := \sup\left\{\int_\Omega u\varphi' : \varphi \in C_c^1(\Omega), |\varphi(\cdot)| \leqslant 1\right\}. \tag{8}$$

We omit the subscript Ω when the domain is clear (or not important) from the context. Then the space BV (Bounded Variation) is defined as $\{u \in L^1(\Omega) : \|Du\|_{TV} < \infty\}$. Equipped with the norm $\|\cdot\|_{BV} := \|\cdot\|_1 + \|\cdot\|_{TV}$, the space BV is Banach. It can be shown that (8) coincides in a subtle sense with the more elementary definition of total variation for functions of one real variable on an interval (a,b), $-\infty < a < b < +\infty$, i.e.,

$$\|Du\|_{TV} := \sup_{\mathcal{P}} \sum_i |u(x_i) - u(x_{i-1})|,$$

where the supremum runs over the set of all finite partitions of (a,b). The coincidence is not quite obvious, for a proof, see for example Theorem 3.27 of Ambrosio et al. (2000).

By Du, we mean the distributional derivative of a L^1_{loc} function u. The spaces $W^{1,p}, 1 \leqslant p \leqslant \infty$ 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 $W^{1,p}, 1 < p \leqslant \infty$ 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 $\mathcal{D} \cap [0,1] \times \mathcal{Q}$ is sequentially 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 space BV, as an generalization of $W^{1,1}$, allows for discontinuous components. In the spirit of Poposition 2, allowing discontinuity might be noteworthy. We will utilize an analogue of Rellich-Kondrachov compactness theorem in space BV.

The intuition for the constraint L in the space $\mathcal{Q}$ is that, the derivatives of $Q(\cdot)$ is not allowed to drastic changes. It implies a large algorithmic design cost. This is similar to the cost on a certain energy functional, like $\int_{\underline{R}}^{\overline{R}} |q'(xr_1)|^2 \mathrm{d}G(r_1)$ in some physics problems.

Simultaneously solving for x^* and $q^*(\cdot)$ 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 (9) and thereby identifying the characteristics of the optimal algorithm.

$$\sup_{q \in \mathcal{Q} \cap D_x} O(q; x) \tag{9}$$

where $D_x := \left\{ q \in L^\infty \left((\underline{R}, \overline{R}) \right) \middle| 0 \leqslant q \leqslant \overline{q}(x) \text{ and } (x,q) \text{ satisfies (6)} \right\}$.[22] 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 2. *Suppose F is supported on $[\underline{a}, \overline{a}]$ with continuous density function $f > 0$, where $-\infty < \underline{a} < \overline{a} < +\infty$. Then, for any $x \in (0,1]$, the objective function $O(x,q)$ is strictly concave with respect to q.*

By Lemma 2, $q^*(x)$ is well-defined as demonstrated in the following proposition.

Proposition 5. *Given $x \in [0,1]$, there exists a unique function $q^*(x)$ optimizing (9).*

5.2 Solving 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 condition. The alternative constraint reads:

$$\underbrace{(\alpha+\gamma)\int_0^{\overline{R}} r_2 \mathrm{d}G(r_2) \int_{\underline{R}}^{\overline{R}} q(xr_1)\mathrm{d}G(r_1)}_{\text{Marginal expected payoff}} + \underbrace{\left[\beta + (\alpha+\gamma)x\int_0^{\overline{R}} r_2 \mathrm{d}G(r_2)\right] \int_{\underline{R}}^{\overline{R}} q'(xr_1) r_1 \mathrm{d}G(r_1)}_{\text{Marginal algorithmic penalty}} = 0, \tag{10}$$

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

where $\int_{\underline{R}_r}^{\overline{R}_r} q'(xr_1)r_1 \mathrm{d}G(r_1)$ is well-defined according to Lebesgue's dominated convergence theorem. Denote (10) in a general form with functional I,

$$\int_{\underline{R}}^{\overline{R}} I(r_1, q, q')\mathrm{d}r_1 = 0.$$

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 $[\underline{R}, \overline{R}]$.

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*

$$\frac{\int_{\underline{R}}^{\overline{R}} \frac{\partial I}{\partial y}(r_1, q^*, q^{*\prime})(q_1 - q^*) + \frac{\partial I}{\partial z}(r_1, q^*, q^{*\prime})(q_1' - q^{*\prime})\mathrm{d}r_1}{\int_{\underline{R}}^{\overline{R}} \frac{\partial I}{\partial y}(r_1, q^*, q^{*\prime})(q_2 - q^*) + \frac{\partial I}{\partial z}(r_1, q^*, q^{*\prime})(q_2' - q^{*\prime})\mathrm{d}r_1} < 0.$$

Assumption 3 is simply satisfied when $(r_1 g'(r_1) + g(r_1))$ is not constant.[23] Assumption 4 effectively assumes there exists an interior solution q^*, i.e., given x, $\exists r_1$ such that $q^*(xr_1) \in (0, \overline{q}(x))$, whilst the corner cases $\{0, \overline{q}(x)\}$ 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 (9) given x.

Theorem 2. *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. $\forall q_1 \in \mathcal{Q}$,*

$$\begin{aligned} 0 \geqslant \int_{\underline{R}}^{\overline{R}} &\big(q_1(xr_1) - q^*(xr_1)\big)\left[k_1(x) - \frac{1}{2}k_2(x)F^{-1}(q^*(xr_1)) + \lambda\frac{\beta}{x} + \lambda(A + \frac{\beta}{x})r_1\frac{g'(r_1)}{g(r_1)}\right]\mathrm{d}G(r_1) \\ &+ \big(q_1(xr_1) - q^*(xr_1)\big)\lambda(A + \frac{\beta}{x})r_1 g(r_1)\Big|_{\underline{R}}^{\overline{R}}. \end{aligned} \tag{11}$$

Now we show how this necessary condition restricts the potential q^* to a specific formula. Consider that the set $U := \{r_1 \in (\underline{R}, \overline{R}) \mid 0 < q^*(xr_1) < \overline{q}(x)\}$ is open, and

[23] Appendix 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.

$C := \{r_1 \in [\underline{R}, \overline{R}] \mid q^*(xr_1) = 0 \text{ or } q^*(xr_1) = \overline{q}(x)\}$ is (relatively) closed. Fix any text function $v \in C_c^\infty(U)$. Then if $|\delta|$ is sufficiently small, $0 \leqslant q_1 := q^* + \delta v \leqslant \overline{q}(x)$, $q_1 \in \mathcal{Q}$. Thus (11) implies[24]

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

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

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

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

$$0 = k_1(x)g(r_1) - \frac{1}{2}k_2(x)F^{-1}(q^*(xr_1))g(r_1) + \lambda\frac{\beta}{x}g(r_1) + \lambda(A + \frac{\beta}{x})r_1g'(r_1) \quad \text{in } U. \tag{12}$$

Notably, when $U = (\underline{R}, \overline{R})$, we can also fix any text function $v \in C_c^\infty(\overline{U})$, where $\overline{U}$ is the closure of U. Then if $|\delta|$ is sufficiently small, $0 \leqslant q_1 := q^* + \delta v \leqslant \overline{q}(x)$ and so $q_1 \in \mathcal{Q}$ thus satisfies (11). Similarly, the inequality must have the equal sign. Together with (12), we obtain

$$0 = \lambda\delta(Ax+\beta)\Big[(\overline{R}g(\overline{R})/x)v(x\overline{R}) - (\underline{R}g(\underline{R})/x)v(x\underline{R})\Big]. \tag{13}$$

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

So far, we draw the conclusion from Theorem 2 that $\exists\lambda \in \mathbb{R}$, s.t.[25]

$$q^*(xr_1) = F\left(\frac{k_1(x) - \lambda\big[\beta/x + (A+\beta/x)r_1g'(r_1)/g(r_1)\big]}{k_2(x)/2}\right). \tag{14}$$

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

$$q^*(xr_1) = F\left(\frac{k_1(x)}{k_2(x)/2}\right). \tag{15}$$

5.3 Non-monotonic 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

[24] Note that since $v \in C_c^\infty(U)$, $v(x\underline{R}) = v(x\overline{R}) = 0$.

[25] The expression $q^* = F(a^*)$ in (14) and (15) implicitly assumes $a^* \in (\underline{a}, \overline{a})$, while the other cases are relatively trivial, corresponding to the algorithm that never/always recommends to each investor.

truncated normal distribution supported on $[-3, 2]$, with $\mu = 0.5$ and $\sigma = 1.5$.[26] In precise, the probability density function of R_t reads

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

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

Figure 3 visualizes the optimal algorithm q^* solved from (14) that ensures the equilibrium $x = 0.4$.[27] 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 $x \leqslant 0.4$. 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.[28]

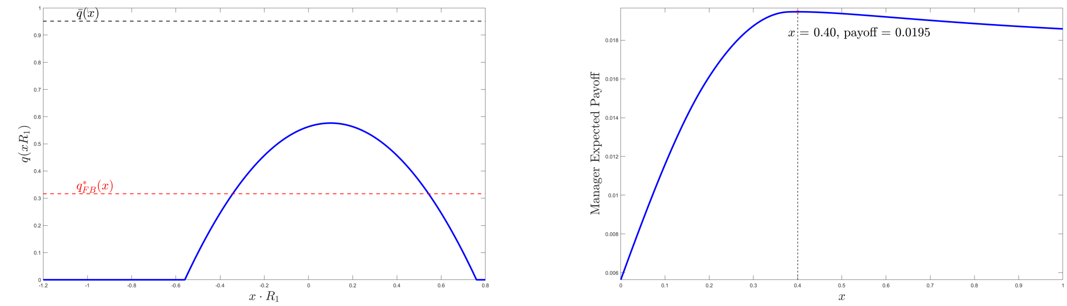
Recall the solving process. Theorem 2 only provides a necessary condition by alternatively satisfying the local incentive constraint–we need to verify that the IC constraint is satisfied. As Panel (b) shows, the algorithm successfully breaks the monotonicity of manager utility: the manager maximizes 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 transfers the investors' risk-aversion to the risk-neutral manager.

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

[26] We 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.

[27] Note 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.

[28] The return interval with $q^* \leqslant q_{FB}(x)$ 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.



(a) Optimal algorithm given $x = 0.4$

(b) Manager's Utility under q^* for different x

Figure 3: Optimal algorithm given $x = 0.4$ and Manager's expected payoff

Notes: Figure 3 illustrates illustrates how the algorithm incentivizes the fund manager to choose $x = 0.4$ rather than $x = 1$ in the continuous case. Panel (a) depicts 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 probability higher than the first-best when returns are moderate, and lower 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 $x = 0.4$. The parameters of contract and the manager's personal benefit are given as $\alpha = \beta = 0.01$ 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 $\sigma = 1.5$.

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*

$$
\begin{aligned}
&x^* \in \arg\max_{x' \in [x^*, 1]} \left\{ (x'A + \beta)[G(x^*\overline{R}/x') - G(x^*\underline{R}/x')] \right\} \text{ and} \\
&x^* \in \arg\max_{x'} \left\{ k_1(x') F\left(\frac{k_1(x')}{k_2(x')/2}\right) - \frac{1}{2}\left(\int_0^{k_1(x')/(k_2(x')/2)} a \mathrm{d}F(a)\right) k_2(x') \right\}.
\end{aligned}
\tag{16}
$$

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. *If $\lambda^* \neq 0$, there exists $r, r' \in [\underline{R}, \overline{R}]$ such that*

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

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 Extended Discussion

6.1 Alternative Information Structure and Algorithm Functionality

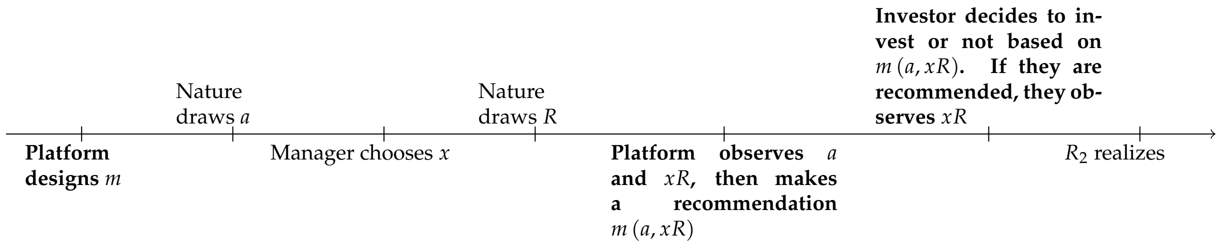
In the baseline model, the roles of the recommendation algorithm are twofold. (i) It processes two information, fund historical returns and investors' risk aversions. (ii) By algorithmic automation, it provides a commitment power which ensures execution even realizes cases with ex-post inefficiency. 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 4, 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 5.

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.

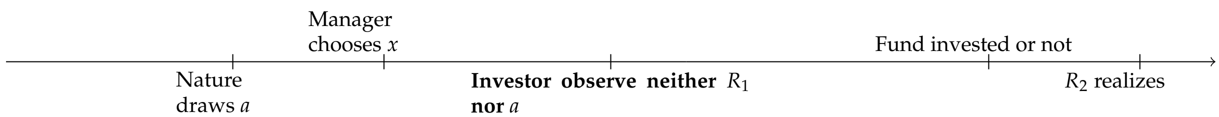
Proposition 7. *Given the contract $\phi(\cdot)$, $x = 1$ is a dominant strategy of the fund manager.*

1. *If $k_1(1) - 1/2\mathbb{E}[a]k_2(1) \geqslant 0$, there exists an equilibrium where all investors invest in the fund and the manager chooses $x^* = 1$.*

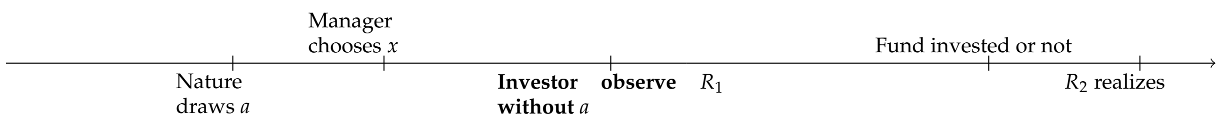
2. *If $k_1(1) - 1/2\mathbb{E}[a]k_2(1) < 0$, there exists an equilibrium where no investor invest in the fund and the manager chooses $x^* = 1$.*



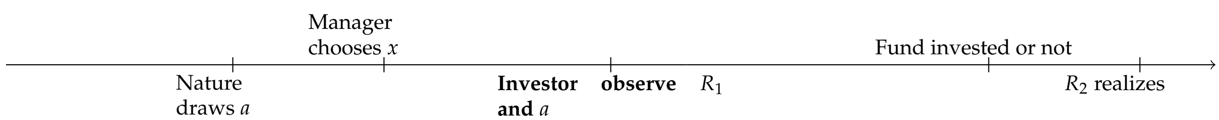
(a) Timeline 1: Baseline Model with Platform and Algorithm



(b) Timeline 2: Model without Information about R_1 and a



(c) Timeline 3: Model without Information about a



(d) Timeline 4: Model with Information about R_1 and a

Figure 4: Alternative Timelines

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 $q_I : [\underline{R}, \overline{R}] \to [0, 1]$.

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

1. *If* $k_1(1) - 1/2\mathbb{E}[a]k_2(1) \geqslant 0$, *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* $(\overline{x}_I, q^*(\cdot))$,

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

where

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

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 infers an x (when the historical return is non-zero) that generates a positive expected return.[29] 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 $\{\underline{R}, \overline{R}\}$, i.e., investors always correctly learn x. Then we have the following implication.

Corollary 2. *Suppose* $supp(R_t) = \{\underline{R}, \overline{R}\}$ *and* $k_1(1) - 1/2\mathbb{E}[a]k_2(1) < 0$, *there is no equilibrium*

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

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 $m_I : [\underline{a}, \overline{a}] \times [\underline{R}, \overline{R}] \to [0, 1]$. For any a, denote the minimum and maximum of $\{x \in [0, 1] : k_1(x) - 1/2ak_2(x) \geqslant 0\}$ as $\underline{x}_I(a)$ and $\overline{x}_I(a)$, respectively, and $\hat{a}(x) = k_1(x)/(k_2(x)/2)$. Proposition 9 analytically solves the equilibrium allocation and the corresponding investment choice.

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

$$
\begin{aligned}
&(x^* A + \beta) \left[p(0) \int_{\underline{a}}^{\hat{a}(x^*)} 1 \mathrm{d}F(a) + p(\overline{R}) \int_{\underline{a}}^{\hat{a}(x^*)} 1 \mathrm{d}F(a) + p(\underline{R}) \int_{\underline{a}}^{\hat{a}(x^*)} 1 \mathrm{d}F(a) \right] \\
&\qquad \geqslant (x A + \beta) \left[p(0) \int_{\underline{a}}^{\hat{a}(x^*)} 1 \mathrm{d}F(a) + p(\overline{R}) \int_{\underline{a}}^{\hat{a}(x)} 1 \mathrm{d}F(a) + p(\underline{R}) \int_{\underline{a}}^{\hat{a}(x)} 1 \mathrm{d}F(a) \right]. \quad (17)
\end{aligned}
$$

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

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

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

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

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

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

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

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.

The deeper intuition is: each investor refuses to invest with ex-post inefficiency, generating no commitment power. Then no one pays for the information rent. In particular, the always-investors would not limit themselves just to force the manager to reduce risk-taking, which expands inclusion and may increase the aggregate expected payoff. This lack of coordination, in a sense, exhibits a curse of shrewdness.

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.

Figure 5 fixes the distribution of risk aversion and the risk return, and compares the distribution of investor welfare under the above settings. The blind investment, as shown in the red dashed line, always results in a full risk taking, and the manager takes away all the investor welfare. Investors on a fund distribution platform, shown by the green solid line, may be over-exposed to risk without the protection of algorithm. The experts in purple never fall into a negative expected payoff. However, only the investors with low risk aversion enters the investment and obtain positive expected payoff (shown in area A), and the population fails to achieve a fairly low x.[30] Ultimately, the blue line shows the baseline with algorithm: by leveraging the whole user base, the algorithm generates a commitment power, achieving an investor-optimal equilibrium $x^* = 0.515$. The piece-wise near $a = 3$ reflects the information rent, whereas the resulting aggregate expected payoff (roughly corresponding to area B) is greater than any other case with significantly expanded financial inclusion than the expert case.

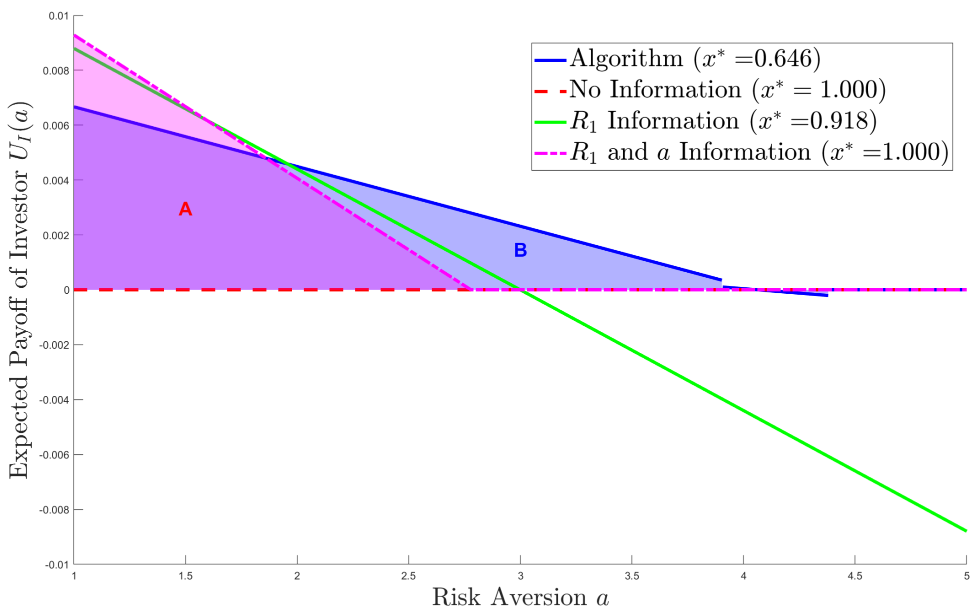


Figure 5: Expected payoff under different risk aversion a and information structures.

Notes: Figure 5 illustrates the distribution of investors' expected utility in equilibrium under four different information structures (ranked by risk aversion a). The blue solid line corresponds to the baseline model, the red dashed line represents the case where the investor has no information (Proposition 7), the green solid line corresponds to the case where the investor can observe R_1 (Proposition 8), and the pink dashed line represents the case where the investor can accurately observe both R_1 and a (Proposition 9). It can be seen that by allowing the investor to access coarse information about their own a, the overall welfare of investors improves, as the Area A is larger than B. 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 above discussion highlights the function of algorithm: it gathers historical information, provides noisy signals about risk aversion, and importantly, 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 con-

[30] In this scenario, the resulting equilibrium allocation $x = 1$, while it is possible to reach an equilibrium with $x < 1$.

sider 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.

6.2 Recommendation Algorithms and Fund Ratings

In practice, asset management and investment advisory are significantly influenced by fund ratings (e.g., Morningstar rating), which are also a mainstream focus in existing literature (e.g., Evans and Sun, 2021; Ben-David et al., 2022; Huang et al., 2020). In practice, recommendation algorithms and fund ratings are not mutually exclusive, as the platform can even publish ratings and deliver personalized recommendation simultaneously. This section discusses the interaction between the algorithms and ratings, and emphasizes the distinctive features of algorithms from two aspects: financial inclusion and personalized information.

Fund ratings are mainly motivated by understanding differences and comparisons between funds: for an investor, which funds are proper to invest? On the other hand, the starting point of recommendation algorithms is the investor heterogeneity: given a fund, who are proper to invest in it? This logic is particularly tractable for connecting mechanism design with financial inclusion. In this context, they share similar functionalities in processing the fund's historical information and for classification.[31] The fey difference is that the ratings are publicly known,[32] while the algorithm delivers private signals according to investors' characteristics, e.g., risk aversion.

To analyze algorithms with the existence of fund ratings, we relax our baseline assumption that investors only know the existence of a fund once they receive a signal. Now, with the existence of publicly known ratings, investors always know the fund and thus can invest even without recommendation.[33] The platform's problem (still along with the discrete setting) can be reformulated as follows: $x \in [0, 1]$,

$$\max_{q:[\underline{R},\overline{R}]\to[0,1]} k_1(x)q(xr_1) - \frac{1}{2}\left(\int_{\underline{a}}^{F^{-1}(q(xr_1))} a\mathrm{d}F(a)\right) k_2(x)\mathrm{d}G(r_1)$$

[31] When inferring an extremely high allocation x, the algorithm effectively classifies it as "high-risk" and thus reduces recommendation.

[32] With 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).

[33] In the present work, we do not model multiple funds, as the competition across funds is beyond our focus. After including two or more funds as mentioned in the setup, our framework allows for fully examining the interaction between algorithms and fund ratings, though it significantly adds technical complexity.

subject to the IC and IR constraints

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

$$k_1(x) - \frac{1}{2} \frac{\int_{\underline{a}}^{F^{-1}(q(xr_1))} a \mathrm{d}F(a)}{q(xr_1)} k_2(x) \geqslant 0, \forall r_1 \in \mathrm{supp}(R_1), \tag{18}$$

$$k_1(x) - \frac{1}{2} \frac{\int_{F^{-1}(q(xr_1))}^{\overline{a}} a \mathrm{d}F(a)}{1 - q(xr_1)} k_2(x) \leqslant 0, \forall r_1 \in \mathrm{supp}(R_1). \tag{19}$$

The additional IR constraint (19) 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 (19) 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 (19) 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

$$\underline{q}(x) := \inf \left\{ q \in [0,1] \middle| k_1(x)(1-q) - \frac{1}{2} \int_{F^{-1}(q)}^{\overline{a}} a \mathrm{d}F(a) k_2(x) \leqslant 0 \right\}.$$

Then the IR constraint (19) is equivalent to $q(xr_1) \geqslant \underline{q}(x)$ for any $r_1 \in [\underline{R}, \overline{R}]$. Proposition 10 indicates how the two IR constraints bind and interact with the equilibrium allocation x.

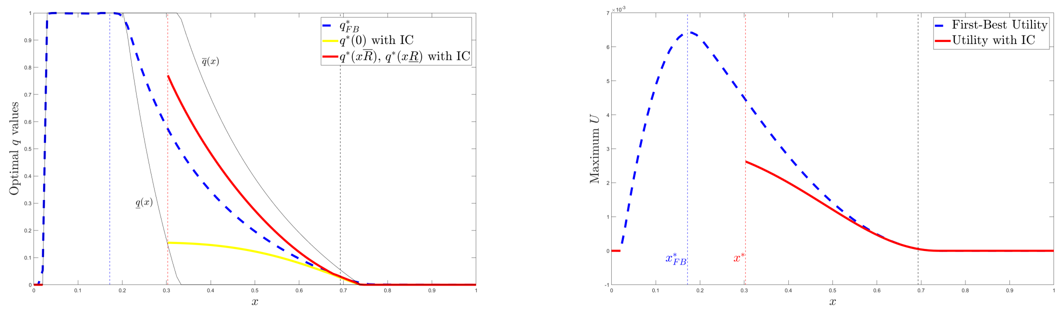
Proposition 10. *Suppose $k_1(x) - 1/2\underline{a}k_2(x) \geqslant 0$.*

1. *If $k_1(x) - \frac{1}{2} \int_{\underline{a}}^{\overline{a}} a \mathrm{d}F(a) k_2(x) < 0$, $\underline{q}(x) = 0$ and $\overline{q}(x) \in (0,1)$.*

2. *If $k_1(x) - \frac{1}{2} \int_{\underline{a}}^{\overline{a}} a \mathrm{d}F(a) k_2(x) > 0$, $\underline{q}(x) \in (0,1)$ and $\overline{q}(x) = 1$.*

Proposition 10 highlights the importance of the population average risk aversion, or the average belief. (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 6 visualizes this case in comparison with the baseline, exactly corresponds to the second scenario in Proposition 10. 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.



(a) $U(x, \hat{q})$ for different x

(b) q^* for different x

Figure 6: Investor's Payoff and $\hat{q}^*(\cdot)$ with constraint $q(xr_1) \geqslant \underline{q}(x)$

Notes: Figure 6 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 suggests a counter-intuitive implication: how could additional public information harm social welfare? Because the algorithm recommendation is relatively less determinative for investors who hold alternative information. Then the lack in driving the user base results in a limited commitment power, and tilts the possible equilibrium allocations 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. Whereas the platform can determine the adoption of fund ratings and recommendation algorithms, the interacted mechanism design suggests an interesting topic for further exploration.

7 Conclusion

We develop a model of recommendation algorithm design in delegated investment. The intermediate platform serves investors who are unaware of their risk aversion levels and aims to mitigate fund managers' moral hazard in over risk-taking, particularly in the cases where the contracts are given. We show that predetermined automatic algorithms can successfully mitigate the principal-agent problem inherent in linear and limited-liability contracts. The core intuition is: the algorithm reshapes the information transmission and further affects the buyer party's entrance, which effectively generates commitment power. More 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 recommendation under ambiguous information and potentially compensates informative signals. This generates an information rent paid by investors that facilitates trading and achieves Pareto improvement. In addition, we provide an approach to general algorithm design problems, discuss inequality issues in algorithm implementation, interaction with contract design, and comparisons with various scenarios of delegated investment.

For instance, this paper focuses on the risk incentives, yet the framework can be extended to designing algorithms for solving other principal-agent problems, e.g., managers' efforts and information acquisition (He and Xiong, 2013; Huang et al., 2020; Buffa et al., 2022). Our analysis also inspires future explorations on many topics, e.g., optimal joint design with contracts, competition with multiple funds, and information design with both public fund ratings and personalized recommendations.

Furthermore, the powerful algorithm necessitates attention to its regulation and mission. 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 tiles to the manager's fund side (e.g., commissions) rather than the user base, the algorithm may create a new principal-agent relationship that is detrimental to investor surplus. This suggests the importance of privacy protection, especially when the platform can know much more about users than the users themselves. In all, our paper uncovers a tip of the algorithm's power in fighting with moral hazard in the context of digital finance. For following research and applications, there is an old saying:

> Whoever fights monsters should see to it that in the process he does not become a monster.
>
> Aphorism 146, *Beyond Good and Evil*, Friedrich W. Nietzsche

References

Adams, Robert A and John JF Fournier, *Sobolev spaces*, 2 ed., Vol. 140 of *Pure and Applied Mathematics*, Elsevier, 2003.

Ambrosio, Luigi, Nicola Fusco, and Diego Pallara, *Functions of bounded variation and free discontinuity problems*, Vol. 46 of *Oxford Mathematical Monographs*, Oxford university press, 2000.

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, Jiacui 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.

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.

Giusti, Enrico, *Minimal surfaces and functions of bounded variation*, Vol. 80 of *Monographs in Mathematics*, Springer Science & Business Media, 1984.

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.

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.

Ravid, Doron, Anne-Katrin Roesler, and Balázs Szentes, "Learning before trading: on the inefficiency of ignoring free information," *Journal of Political Economy*, 2022, *130* (2), 346–387.

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.

_ , Youchang Wu, and Josef Zechner, "Intermediated investment management," *The Journal of Finance*, 2011, *66* (3), 947–980.

Sun, Jian, "Algorithmic transparency," *Unpublished Working Paper*, 2024.

Szalay, Dezsö, "The economics of clear advice and extreme options," *The Review of Economic Studies*, 2005, *72* (4), 1173–1198.

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 algorithm m in an equilibrium. For any fund with historical r_{p1}, let $q_m(r_{p1}) := \int_{\underline{a}}^{\overline{a}} m(a, r_{p1})\mathrm{d}F(a)$ 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:

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

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

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

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 $rp1$.

A.2 Proof of Proposition 2

(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),

$$(Ax' + \beta) \int_{\underline{R}}^{\overline{R}} \hat{q}(x'r_1) \mathrm{d}G(r_1) \leqslant (Ax' + \beta) \int_{\underline{R}}^{\overline{R}} q^*(x'r_1) \mathrm{d}G(r_1).$$

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

$$(Ax' + \beta) \int_{\underline{R}}^{\overline{R}} q^*(x'r_1) \mathrm{d}G(r_1) \leqslant (Ax^* + \beta) \int_{\underline{R}}^{\overline{R}} q^*(x^*r_1) \mathrm{d}G(r_1) = (Ax^* + \beta) \int_{\underline{R}}^{\overline{R}} \hat{q}(x^*r_1) \mathrm{d}G(r_1).$$

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

A.3 Proof of Proposition 3

Under this $x \in [0, 1]$, the designer's problem can be rewritten by

$$\max_{q:[\underline{R},\overline{R}]\to[0,1]} k_1(x) q(xr_1) - \frac{1}{2} \left(\int_{\underline{a}}^{F^{-1}(q(xr_1))} a \mathrm{d}F(a) \right) k_2(x) \mathrm{d}G(r_1)$$

subject to the IC and IR constraint

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

$$q(x) \leqslant \overline{q}(x). \tag{A2}$$

Since x takes continuous values on [0,1], all values of q on $[\underline{R}, \overline{R}]$ 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), 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 $(x, q(x))$.

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

As a result, if there exists a solution $q^*(x)$, then it can be found in the set $\{\hat{q}(x)\}$.

Under $\hat{q}$, the objective function is

$$\max_{q(x\underline{R}),q(x\overline{R}),q(0)} k_1(x)\left[\hat{q}(x\underline{R})p(\underline{R})+\hat{q}(0)p(0)+\hat{q}(x\overline{R})p(\overline{R})\right]$$
$$-\frac{1}{2}k_2(x)\left[\left(\int_{\underline{a}}^{F^{-1}(\hat{q}(x\underline{R}))} a\mathrm{d}F(a)\right)\hat{q}(x\underline{R})+\left(\int_{\underline{a}}^{F^{-1}(\hat{q}(0))} a\mathrm{d}F(a)\right)\hat{q}(0)+\left(\int_{\underline{a}}^{F^{-1}(\hat{q}(x\overline{R}))} a\mathrm{d}F(a)\right)\hat{q}(x\overline{R})\right].$$

Without the IC constraint, for any x, the optimal algorithm should recommend all investors whose payoff is non-negative under x. Then we have $q_{FB}(x) = k_1(x)/(k_2(x)/2)$ as the first-best algorithm without the IC constraint and x^*_{FB} satisfies

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

In the following, we call the equilibrium characterized by $(x^*_{FB}, q_{FB}(x^*_{FB}))$ is the *First-Best Equilibrium*.

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

$$(Ax+\beta)\left[\hat{q}(x\underline{R})p(\underline{R})+\hat{q}(0)p(0)+\hat{q}(x\overline{R})p(\overline{R})\right] \geqslant \max_{x'}(Ax'+\beta)\hat{q}(0)p(0) = (A+\beta)\hat{q}(0)p(0).$$

The modification on the left side of the IC constraint arises from the scenario where the distribution of risky returns is represented by a three-point discrete model. In this model, the manager's deviation from x to x' remains undetectable only when the historic return equals zero. When the historic return on the risky asset r_1 is strictly positive or negative, a clear inequality $xr_1 \neq x'r_1$ is established. Consequently, given the algorithm, the probability of being recommended is given by $\hat{q}(0)$ and the manager will deviate toward $x' = 1$ only to ensure that the expected return is maximized when he is recommended. As a result, the incentive compatibility constraint needs only to ensure that the fund manager does not prefer $x' = 1$. The case of continuous distributions is more complicated because small deviations in the manager's actions are not easily detected by the algorithm.

Given x, by Kuhn-Tucker conditions, we have the following Lagrangian:

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

$$+\eta_{\underline{R},1}\hat{q}(x\underline{R})+\eta_{\underline{R},2}(\overline{q}(x)-\hat{q}(x\underline{R}))+\eta_{\overline{R},1}\hat{q}(x\overline{R})+\eta_{\overline{R},2}(\overline{q}(x)-\hat{q}(x\overline{R}))+\eta_{0,1}\hat{q}(0)+\eta_{0,2}(\overline{q}(x)-\hat{q}(0)), \tag{A5}$$

where $\lambda \geqslant 0$ and $\eta_{r,1} \geqslant 0$ are the multipliers from IC and IR constraints, and $\eta_{r,2} \geqslant 0$ 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 (recalling that a follows an uniform distribution),

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

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

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

the IC constraint is satisfied. Additionally, since for a given x^*, $q_{FB}(x^*)$ maximizes the objective function and $q_{FB}(x^*) \leqslant \underline{q}(x^*)$, then if $x^* \geqslant \underline{x}$, $\hat{q}^*(x^*\underline{R}) = \hat{q}^*(0) = \hat{q}^*(x^*\overline{R}) = q_{FB}(x^*)$ satisfies all constraints and maximizes the objective function, making it a solution.

If $x^* < \underline{x}$, (A6) does not hold. Therefore, $\lambda > 0$, which further implies $\min\{\hat{q}^*(x^*\underline{R}), \hat{q}^*(x^*\overline{R})\} \leqslant q_{FB}(x^*) \leqslant \hat{q}^*(0)$.

A.4 Proof of Proposition 4

We primarily discuss the proof of the interior solution.

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

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

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

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

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

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

$$k_1'(x^*)q^E(x^*) - \frac{1}{2}k_2'(x^*)\left[p(0)\int_{\underline{a}}^{F^{-1}(\hat{q}^*(0))} a\mathrm{d}F(a) + \sum_{r \in \{\underline{R}, \overline{R}\}} p(r)\int_{F\underline{a}}^{F^{-1}(\hat{q}^*(xr))} a\mathrm{d}F(a)\right] + \lambda A q^E = 0$$
$$\Rightarrow k_1'(x^*)q^E(x^*) - \frac{1}{2}k_2'(x^*)\left[p(0)\int_{\underline{a}}^{F^{-1}(\hat{q}^*(0))} a\mathrm{d}F(a) + \sum_{r \in \{\underline{R}, \overline{R}\}} p(r)\int_{F\underline{a}}^{F^{-1}(\hat{q}^*(xr))} a\mathrm{d}F(a)\right] < 0.$$

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

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

we have

$$k_1'(\tilde{x})q^E(x^*) - \frac{1}{2}k_2'(\tilde{x})\left[p(0)\int_{\underline{a}}^{F^{-1}(\hat{q}^*(0))} a\mathrm{d}F(a) + \sum_{r \in \{\underline{R}, \overline{R}\}} p(r)\int_{F\underline{a}}^{F^{-1}(\hat{q}^*(xr))} a\mathrm{d}F(a)\right] < 0.$$

A.5 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 $\lim_{n\to\infty} x_n \to b$. By Lipschitzness of u, $\{u(x_n)\}_{n=1}^{\infty}$ is a Cauchy sequence, and thus $\lim_{n\to\infty} u(x_n)$ exists. The uniqueness of this limit among all such sequences then follows again from the Lipschitzness of u, i.e., $\lim_{x\to b} u(x)$ exists. Further, we have

$$|u(x_0) - u(b)| = \lim_{n\to\infty} |u(x_0) - u(x_n)| \leqslant \lim_{n\to\infty} L|x_0 - x_n| = L|x_0 - x|$$

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 $\mathcal{Q}' = \left\{q \in C_b\left(\left[\underline{R}, \overline{R}\right]\right) : 0 \leqslant q(\cdot) \leqslant 1 \text{ and } \sup_{x\neq y} \frac{|q(x)-q(y)|}{|x-y|} \leqslant L\right\}$. For all $\varepsilon > 0$, choose $\delta = \varepsilon/L$. Then for all $q \in \mathcal{Q}'$, $x, y \in \left[\underline{R}, \overline{R}\right]$ 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 $C\left(\left[\underline{R}, \overline{R}\right]\right)$. By the unique extension, we can identify $\mathcal{Q}$ with $\mathcal{Q}'$, so $\mathcal{Q}$ is also precompact in $C_b\left(\left(\underline{R}, \overline{R}\right)\right)$. Let $\{q_n\}_{n=1}^{\infty}$ be a sequence in $\mathcal{Q}$ and $q_n \to q$ under the sup norm. For all $x, y \in \left(\underline{R}, \overline{R}\right)$, we have

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

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 $(x_n, q_n) \to (x, q)$ for $q \in \mathcal{Q}$. By definition, for all $x' \in [0, 1]$ and each $n \in \{1, 2, \ldots\}$, we have

$$(Ax_n + \beta) \int_{\underline{R}}^{\overline{R}} q_n(x_n r_1)\, dG(r_1) \geqslant \left(Ax' + \beta\right) \int_{\underline{R}}^{\overline{R}} q_n\left(x' r_1\right) dG(r_1). \tag{A8}$$

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

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

On the other hand,

$$\begin{aligned}\left|\int_{\underline{R}}^{\overline{R}} q_n(x_n r_1)\, dG(r_1) - \int_{\underline{R}}^{\overline{R}} q(x r_1)\, dG(r_1)\right| \\ &\leqslant \int_{\underline{R}}^{\overline{R}} |q_n(x_n r_1) - q(x r_1)|\, dG(r_1) \\ &\leqslant \int_{\underline{R}}^{\overline{R}} (|q_n(x_n r_1) - q_n(x r_1)| + |q_n(x r_1) - q(x r_1)|)\, dG(r_1) \\ &\leqslant \int_{\underline{R}}^{\overline{R}} \left(L\,|x_n - x|\, r_1 + \sup_{y\in(\underline{R},\overline{R})} |q_n(y) - q(y)| \right) dG(r_1) \to 0.\end{aligned} \tag{A9}$$

The convergence of the right hand side of the last inequality follows again from the dominated convergence theorem. Therefore, as $n \to \infty$, the left hand side of (A8) converges and we get

$$(Ax+\beta)\int_{\underline{R}}^{\overline{R}} q(x r_1)\, dG(r_1) \geqslant \left(Ax'+\beta\right)\int_{\underline{R}}^{\overline{R}} q\left(x' r_1\right) dG(r_1).$$

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

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

By (A9) and the continuity of $k_1(x)$ and $k_2(x)$, we only need to consider the term

$$\int_{\underline{R}}^{\overline{R}} \int_{\underline{a}}^{F^{-1}(q(x r_1))} a\, \mathrm{d}F(a)\mathrm{d}G(r_1).$$

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

$$\left|\int_{\underline{a}}^{F^{-1}(q_n(x_n r_1))} af(a)\, da - \int_{\underline{a}}^{F^{-1}(q(x r_1))} af(a)\, da\right| \leqslant K\left|F^{-1}(q_n(x_n r_1)) - F^{-1}(q(x r_1))\right|$$

for some constant K. Since F is strictly increasing and continuous, F^{-1} is strictly increasing and continuous too. Similar to (A9), the right hand side of the above inequality converges to 0 as $n \to \infty$. Then by dominated convergence theorem, we have

$$\int_{\underline{R}}^{\overline{R}} \left(\int_{\underline{a}}^{F^{-1}(q_n(x_n r_1))} af(a)\, da\right) dG(r_1) \to \int_{\underline{R}}^{\overline{R}} \left(\int_{\underline{a}}^{F^{-1}(q(x r_1))} af(a)\, da\right) dG(r_1).$$

Here we use the obvious fact that $\left|\int_{\underline{a}}^{F^{-1}(q_n(x_n r_1))} af(a)\, da\right| \leqslant \mathbb{E}[\boldsymbol{a}] < \infty$.

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

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

For all $q \in \mathcal{Q}$, $\|q\|_{1,p} = \|q\|_p + \|Dq\|_p \leqslant \left(|\underline{R}| + |\overline{R}|\right) + L$, i.e., $\mathcal{Q}$ is a bounded in norm. By Banach-Alaoglu-Bourbaki theorem and the fact that $W^{1,p}\left(\left(\underline{R}, \overline{R}\right)\right)$ is reflective 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 $\|q_n - q\|_{1,p} \to 0$ as $n \to \infty$. Since $\|\cdot\|_{1,p}$ dominates $\|\cdot\|_p$, we have $0 \leqslant q\left(\cdot\right) \leqslant 1$ a.e. with respect to Lebesgue measure. We also have that for all $\varepsilon > 0$, $\|Dq\|_p \leqslant \|Dq_n\|_p + \|Dq_n - Dq\|_p \leqslant \|Dq_n\|_p + \|q_n - q\|_{1,p} \leqslant L + \varepsilon$, when $n\left(\varepsilon\right)$ is chosen large enough. By the arbitrariness of ε, $\|Dq\| \leqslant L$, 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 Thereom 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 $q \in W^{1,p}\left(\left(\underline{R}, \overline{R}\right)\right)$ with its $C^{0,1-\frac{1}{p}}\left(\left[\underline{R}, \overline{R}\right]\right)$ version. This identification is possible due to Morrey inequality and the regularity of the boundary of $\left(\underline{R}, \overline{R}\right)$, see for example Theorem 5 in subsection 5.8.4 of Evans (2010). To be more specific[34], we show that $\{\underline{R}, \overline{R}\}$ is a C^1 (and thus Lipschitz) boundary. Consider $\left\{B\left(\underline{R}, \rho\right), B\left(\overline{R}, \rho\right)\right\}$, where $B\left(x_0, \mathrm{r}\right)$ means the open ball centered at x_0 with radius r, and we may take $\rho = \frac{1}{4}\left(|\underline{R}| + |\overline{R}|\right)$. Let $\mathbb{R}^0 = \{0\}$, then $f_1 \equiv \underline{R}$, $f_2 \equiv \overline{R}$ are C^1 functions on $\mathbb{R}^0$. Obviously, $\left(\underline{R}, \overline{R}\right) \cap B\left(\underline{R}, \rho\right) = \{r \in B\left(\underline{R}, \rho\right) : r > f_1\}$ and $\left(\underline{R}, \overline{R}\right) \cap B\left(\overline{R}, \rho\right) = \left\{r \in B\left(\overline{R}, \rho\right) : r < f_2\right\}$, i.e., $\{\underline{R}, \overline{R}\}$ is the C^1 boundary of $\left(\underline{R}, \overline{R}\right)$.

Now, by the compact embedding $W^{1,p} \hookrightarrow C^{0,1-\frac{1}{p}-\epsilon}$ and the discussion in the next step,

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

we can modify (A9) to

$$\left|\int_{\underline{R}}^{\overline{R}} q_n\left(x_n r_1\right) dG\left(r_1\right)-\int_{\underline{R}}^{\overline{R}} q\left(x r_1\right) dG\left(r_1\right)\right|$$
$$\leqslant \int_{\underline{R}}^{\overline{R}}\left(\left|q_n\left(x_n r_1\right)-q_n\left(x r_1\right)\right|+\left|q_n\left(x r_1\right)-q\left(x r_1\right)\right|\right) dG\left(r_1\right)$$
$$\leqslant \int_{\underline{R}}^{\overline{R}}\left(L'\left(\left|x_n r_1-x r_1\right|\right)^{1-\frac{1}{p}-\epsilon}+\sup_{y \in\left(\underline{R}, \overline{R}\right)}\left|q_n(y)-q(y)\right|\right) dG\left(r_1\right) \rightarrow 0, \tag{A10}$$

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–Ŝmulian theorem. The continuity of $\bar{q}(x)$ again guarantees that $q(\cdot) \leqslant \bar{q}(x)$.

Step 3: The objective function $O : D \cap ([0,1] \times \mathcal{Q}) \rightarrow \mathbb{R}$ is weakly sequentially continuous, i.e., $(x_n, q_n) \rightharpoonup (x, q)$ 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 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 $\lim_{k\to\infty} \mathcal{L}_1(q_{n_k}) = \mathcal{L}_1(q') \neq \mathcal{L}_1(q)$ or $\lim_{k\to\infty} \mathcal{L}_2(q_{n_k}) = \mathcal{L}_2(q') \neq \mathcal{L}_2(q)$, a contradiction. Therefore $q = q'$, i.e., $q_n \rightarrow q$ in $C^{0,1-\frac{1}{p}-\epsilon}$. Now by (A10) and a discussions similiar 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

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

part 3 The functions of bounded variation on $\left(\underline{R}, \overline{R}\right)$ may not be continuous, since monotonic functions are all of bounded variations. The main difficulty here is that we do not directly have continuity arguments like (A9) or (A10).

Step 1: Reestimating the difference.

Let $\{(x_n, q_n)\}_{n=1}^{\infty}$ be a convergent sequence in $[0,1] \times \mathcal{Q}$, where q_n converges in the L^1 sense to q. From *Step 2* and *Step 3* of both **part 1** and **part 2** we know it is crucial to estimate the integration of $|q_n(x_n r_1) - q(xr_1)|$ on $(\underline{R}, \overline{R})$. Note that we also have

$$\int_{\underline{R}}^{\overline{R}} |q_n(x_n r_1) - q(xr_1)| dG(r_1) \leqslant \underbrace{\int_{\underline{R}}^{\overline{R}} |q_n(x_n r_1) - q(x_n r_1)| dG(r_1)}_{\Delta_{1,n}} + \underbrace{\int_{\underline{R}}^{\overline{R}} |q(x_n r_1) - q(xr_1)| dG(r_1)}_{\Delta_{2,n}}.$$

For the $\Delta_{1,n}$ term, suppose $x_n \neq 0$, then

$$\int_{\underline{R}}^{\overline{R}} |q_n(x_n r_1) - q(x_n r_1)| dG(r_1) = \frac{1}{x_n} \int_{x_n \underline{R}}^{x_n \overline{R}} |q_n(y) - q(y)| g\left(\frac{y}{x_n}\right) dy \leqslant \frac{M}{x_n} \|q_n - q\|_1,$$

for a constant M. Therefore, for $\lim_{n\to\infty} x_n \neq 0$, $\Delta_{1,n} \to 0$ as $n \to \infty$. Note that without loss of generality, we can assume $\lim_{n\to\infty} x_n \neq 0$, since if $\beta \geqslant 0$, $O(0, q) \leqslant 0$ and achieve its maximum value when $q \equiv 0$. For the $\Delta_{2,n}$ term, consider the $C_c\left((\underline{R}, \overline{R})\right)$ approximation of $L^1\left((\underline{R}, \overline{R})\right)$ functions. Since C_c is dense in L^1, such kind of approximation is possible. For all $\varepsilon > 0$, there exists $p(\cdot)$ such that $|p - q| < \varepsilon$, so we have

$$\begin{aligned}\Delta_{2,n} &\leqslant \int_{\underline{R}}^{\overline{R}} \left(|p(x_n r_1) - q(x_n r_1)| + |p(xr_1) - q(xr_1)| + |p(x_n r_1) - p(xr_1)|\right) dG(r_1) \\ &\leqslant M_1 \varepsilon + M_2(p) |x_n - x|,\end{aligned}$$

where M_1 and $M_2(q)$ are some constants, M_1 only depends on G and x. By the arbitrarness of ε, we see $\Delta_{2,n} \to 0$ as $n \to \infty$.

Step 2: Sequentially compactness and sequentially continuity in a weak sense.

We first state a fundamental semi-continuity of the total variation. Let $\Omega \subset \mathbb{R}^n$ be an open set and $\{f_i\}_{i=1}^{\infty}$ a sequence in $BV(\Omega)$ that converges in $L^1_{loc}(\Omega)$ to a function f, then $\|Df\|_{TV} \leqslant \liminf_{i\to\infty} \|Df_i\|_{TV}$ (see for example Theorem 1.9 in Giusti (1984)). Let $\{(x_n, q_n)\}_{n=1}^{\infty}$ be a sequence in $D \cap ([0,1] \times \mathcal{Q})$ such that $\lim_{n\to\infty}(x_n, q_n) = (x, q)$, where q_n converges in the L^1 sense to $q \in L^1$. By the semi-continuity of $\|D \cdot\|_{TV}$, we have $\|Dq\|_{TV} \leqslant \liminf_{n\to\infty} \|Dq_n\|_{TV}$ and thus $\|Dq\|_{TV} \leqslant L$. Since $\|q_n - q\|_1 \to 0$, we also have that $0 \leqslant q \leqslant 1$, i.e., $q \in \mathcal{Q}$.

By the above discussion and *Step 1* of **part 3**, we only need to find a convergent subsequence $\{(x_{n_k}, q_{n_k})\}_{k=1}^{\infty}$ of a maximizing sequence $\{(x_n, q_n)\}_{n=1}^{\infty}$ in $D \cap ([0,1] \times \mathcal{Q})$ (recall the proof idea in **part 2**), where q_{n_k} converges to some q in the L^1 sense. Fortunately, there is a Rellich-Kondrachov type theorem for the space BV (see for example Theorem 1.19 of Giusti (1984)). The theorem states that a set of functions uniformly bounded in $\|\cdot\|_{BV}$ is precompact in L^1. Since $\|q_n - q\|_1 \to 0$, $q_n \leqslant \bar{q}(x_n)$, $x_n \to x$ and the continuity of $\bar{q}(\cdot)$ implies $q(\cdot) \leqslant \bar{q}(x)$ a.e., the proof is finished.

A.6 Proof of Lemma 2

Consider the auxiliary function function

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

By direct calculation, we have

$$\frac{dI^2(q;x)}{dx^2} = -\frac{k_2(x)}{2f(F^{-1}(q))} < 0, \quad \forall q \in [0,1]. \tag{A11}$$

The objective function $O(x,q)$ can be rewritten as $O(x,q) = \mathbb{E}[I(q(x\boldsymbol{r}_1);x)]$. By (A11), $I(\cdot;x)$ 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

$$\begin{aligned} O(x, \lambda q_1 + (1-\lambda) q_2) &= \mathbb{E}[I(\lambda q_1(x\boldsymbol{r}_1) + (1-\lambda) q_2(x\boldsymbol{r}_1))] \\ &> \mathbb{E}[\lambda I(q_1(x\boldsymbol{r}_1)) + (1-\lambda) q_2(x\boldsymbol{r}_1)], \quad \forall x \in (0,1], \forall \lambda \in (0,1), \end{aligned}$$

since by assumption $g > 0$.

A.7 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 $q_n \to q$. By definition, we have

$$\begin{aligned} &\lim_{n\to\infty} A\int_{\underline{R}}^{\overline{R}} q_n(xr_1)\mathrm{d}G(r_1) + (Ax+\beta)\int_{\underline{R}}^{\overline{R}} q_n'(xr_1)r_1\mathrm{d}G(r_1) = 0 \\ &\Rightarrow A\int_{\underline{R}}^{\overline{R}} q(xr_1)\mathrm{d}G(r_1) + (Ax+\beta)\int_{\underline{R}}^{\overline{R}} q'(xr_1)r_1\mathrm{d}G(r_1) = 0. \end{aligned}$$

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

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

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

$$\int_{\underline{R}}^{\overline{R}} u_I[\lambda q_1(xr_1) + (1-\lambda)q_2(xr_1)]\mathrm{d}G(r_1) > \int_{\underline{R}}^{\overline{R}} \lambda u_I[q(xr_1)] + (1-\lambda)u_I[q_2(xr_1)]\mathrm{d}G(r_1)$$
$$\Rightarrow \int_{\underline{R}}^{\overline{R}} u_I[(\lambda q_1 + (1-\lambda)q_2)(xr_1)]\mathrm{d}G(r_1) > \lambda \int_{\underline{R}}^{\overline{R}} u_I[q(xr_1)]\mathrm{d}G(r_1) + (1-\lambda)\int_{\underline{R}}^{\overline{R}} u_I[q_2(xr_1)]\mathrm{d}G(r_1).$$

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.8 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

$$\frac{\int_{\underline{R}}^{\overline{R}} \frac{\partial I}{\partial y}(r_1, q^*, q^{*\prime})(q_1 - q^*) + \frac{\partial I}{\partial z}(r_1, q^*, q^{*\prime})(q_1' - q^{*\prime})\mathrm{d}r_1}{\int_{\underline{R}}^{\overline{R}} \frac{\partial I}{\partial y}(r_1, q^*, q^{*\prime})(q_2 - q^*) + \frac{\partial I}{\partial z}(r_1, q^*, q^{*\prime})(q_2' - q^{*\prime})\mathrm{d}r_1} < 0$$
$$\Leftrightarrow \frac{\int_{\underline{R}}^{\overline{R}} \left[\frac{\partial I}{\partial y}(r_1, q^*, q^{*\prime}) - \frac{d}{dr_1}\left(\frac{\partial I}{\partial z}(r_1, q^*, q^{*\prime})\right)\right](q_1 - q^*)\mathrm{d}r_1 + (Ax+\beta)r_1(q_1 - q^*)|_{\underline{R}}^{\overline{R}}}{\int_{\underline{R}}^{\overline{R}} \left[\frac{\partial I}{\partial y}(r_1, q^*, q^{*\prime}) - \frac{d}{dr_1}\left(\frac{\partial I}{\partial z}(r_1, q^*, q^{*\prime})\right)\right](q_2 - q^*)\mathrm{d}r_1 + (Ax+\beta)r_1(q_2 - q^*)|_{\underline{R}}^{\overline{R}}} < 0$$

and

$$\int_{\underline{R}}^{\overline{R}} \left[\frac{\partial I}{\partial y}(r_1, q^*, q^{*\prime}) - \frac{d}{dr_1}\left(\frac{\partial I}{\partial z}(r_1, q^*, q^{*\prime})\right)\right](q_2 - q^*)\mathrm{d}r_1 + (Ax+\beta)r_1(q_2 - q^*)|_{\underline{R}}^{\overline{R}} \neq 0. \tag{A12}$$

(A12) is possible because of Assumption 3. Then for each $0 \leqslant \tau \leqslant 1$ and $0 \leqslant \delta \leqslant 1$,

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

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

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

Clearly,

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

In addition, i is C^1 and

$$\frac{\partial i}{\partial \tau}(\tau,\delta) = (1-\delta)\int_{\underline{R}}^{\overline{R}} \frac{\partial I}{\partial y}(r_1,\tilde{q},\tilde{q}')(q_2-q^*) + \frac{\partial I}{\partial z}(r_1,\tilde{q},\tilde{q}')(q_2'-q^{*\prime})\mathrm{d}r_1, \tag{A13}$$

$$\frac{\partial i}{\partial \delta}(\tau,\delta) = \int_{\underline{R}}^{\overline{R}} \frac{\partial I}{\partial y}(r_1,\tilde{q},\tilde{q}')[q_1-(q^*+\tau(q_2-q^*))] + \frac{\partial I}{\partial z}(r_1,\tilde{q},\tilde{q}')[q_1'-(q^{*\prime}+\tau(q_2'-q^{*\prime}))]\mathrm{d}r_1. \tag{A14}$$

Consequently (A12) implies that

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

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

$$i(\phi(\delta),\delta) = 0 \tag{A15}$$

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 (A13) and (A14) yield

$$\phi'(0) = -\frac{\frac{\partial i}{\partial \delta}(\phi(0),0)}{\frac{\partial i}{\partial \tau}(\phi(0),0)} = -\frac{\int_{\underline{R}}^{\overline{R}} \frac{\partial I}{\partial y}(r_1,q^*,q^{*\prime})(q_1-q^*) + \frac{\partial I}{\partial z}(r_1,q^*,q^{*\prime})(q_1'-q^{*\prime})\mathrm{d}r_1}{(1-\delta)\int_{\underline{R}}^{\overline{R}} \frac{\partial I}{\partial y}(r_1,q^*,q^{*\prime})(q_2-q^*) + \frac{\partial I}{\partial z}(r_1,q^*,q^{*\prime})(q_2'-q^{*\prime})\mathrm{d}r_1}. \tag{A16}$$

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

Now set

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

and write

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

Since (A15), we see that $q^* + \tilde{q}(\delta) \in D_x \cap \mathcal{Q}$.

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

$$0 \geqslant o'(0) = \int_{\underline{R}}^{\overline{R}} k_1(x)\big[\phi'(0)(q_2 - q^*) + q_1 - q^*\big] - \frac{1}{2}k_2(x)F^{-1}(q^*)\big[\phi'(0)(q_2 - q^*) + q_1 - q^*\big]\mathrm{d}G(r_1). \tag{A17}$$

Define

$$\lambda := \frac{\int_{\underline{R}}^{\overline{R}} k_1(x)(q_2 - q^*) - \frac{1}{2}k_2(x)F^{-1}(q^*)(q_2 - q^*)\mathrm{d}G(r_1)}{(1-\delta)\int_{\underline{R}}^{\overline{R}} \frac{\partial I}{\partial y}(r_1, q^*, q^{*\prime})(q_2 - q^*) + \frac{\partial I}{\partial z}(r_1, q^*, q^{*\prime})(q_2' - q^{*\prime})\mathrm{d}r_1}, \tag{A18}$$

then we see that

$$\begin{aligned} &\phi'(0)\int_{\underline{R}}^{\overline{R}} k_1(x)(q_2 - q^*) - \frac{1}{2}k_2(x)F^{-1}(q^*)(q_2 - q^*)\mathrm{d}G(r_1) \\ &\qquad = -\lambda \int_{\underline{R}}^{\overline{R}} \frac{\partial I}{\partial y}(r_1, q^*, q^{*\prime})(q_1 - q^*) + \frac{\partial I}{\partial z}(r_1, q^*, q^{*\prime})(q_1' - q^{*\prime})\mathrm{d}r_1 \end{aligned}$$

and plugging this into (A17), we have

$$\begin{aligned} 0 &\geqslant \int_{\underline{R}}^{\overline{R}} (q_1 - q^*)\left[k_1(x) - \frac{1}{2}k_2(x)F^{-1}(q^*)\right]\mathrm{d}G(r_1) \\ &\quad - \lambda\int_{\underline{R}}^{\overline{R}} \frac{\partial I}{\partial y}(r_1, q^*, q^{*\prime})(q_1 - q^*) + \frac{\partial I}{\partial z}(r_1, q^*, q^{*\prime})(q_1' - q^{*\prime})\mathrm{d}r_1 \\ &= \int_{\underline{R}}^{\overline{R}} (q_1 - q^*)\left[k_1(x) - \frac{1}{2}k_2(x)F^{-1}(q^*)\right]\mathrm{d}G(r_1) - \lambda\int_{\underline{R}}^{\overline{R}} A(q_1 - q^*) + (Ax + \beta)r_1(q_1' - q^{*\prime})\mathrm{d}G(r_1) \\ &= \int_{\underline{R}}^{\overline{R}} (q_1 - q^*)\left[k_1(x) - \frac{1}{2}k_2(x)F^{-1}(q^*)\right]\mathrm{d}G(r_1) - \lambda\left[Ag(r_1) - (Ax + \beta)(g(r_1) + r_1 g'(r_1))/x\right]\mathrm{d}r_1 \end{aligned}$$

$$
\begin{aligned}
&+ \lambda(Ax+\beta)r_1 g(r_1)/x[q_1(xr_1) - q^*(xr_1)]\Big|_{\underline{R}}^{\overline{R}} \\
=& \int_{\underline{R}}^{\overline{R}} (q_1 - q^*) \left[k_1(x) - \frac{1}{2}k_2(x)F^{-1}(q^*) + \lambda\beta/x + \lambda(A+\beta/x)r_1 g'(r_1)/g(r_1)\right] \mathrm{d}G(r_1) \\
&+ \lambda(Ax+\beta)r_1 g(r_1)/x\,[q_1(xr_1) - q^*(xr_1)]\Big|_{\underline{R}}^{\overline{R}}.
\end{aligned}
$$

Example 1. Consider that the risk return R_t follows a uniform distribution $U[\underline{R}, \overline{R}]$, 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:

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

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

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

Note that

$$
\begin{aligned}
&\arg\max_{x' \leqslant x^*} (x'A+\beta)\hat{q}\Big[G(x'\overline{R}/x') - G(x'\underline{R}/x') \Big] = x^*; \\
&\arg\max_{x' \geqslant x^*} (x'A+\beta)\hat{q}\Big[G(x^*\overline{R}/x') - G(x^*\underline{R}/x') \Big] = \arg\max_{x' \geqslant x^*} x^*(A+\beta/x')\hat{q} = x^*,
\end{aligned}
$$

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.9 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 (14), there exists $r, r' \in [\underline{R}, \overline{R}]$ 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

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

A.10 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, $k_1(1) - 1/2\mathbb{E}[a]k_2(1) \geqslant 0$, 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.

A.11 Proof of Proposition 8

If $k_1(1) - 1/2\mathbb{E}[a]k_2(1) \geqslant 0$, 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 $\{x \in [0,1] : k_1(x) - 1/2\mathbb{E}[a]k_2(x) \geqslant 0\}$ as $\underline{x}_I$ and $\overline{x}_I \in (0,1)$, 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 $r \in (\overline{x}_I\underline{R}, \underline{x}_I\underline{R}) \cup (\underline{x}_I\overline{R}, \overline{x}_I\overline{R})$, investors always should invest in the fund, that is $q_I(r) = 1$. Similarly, given $r \in [\underline{R}, \overline{x}_I\underline{R}) \cup (\underline{x}_I\underline{R}, 0) \cup (0, \underline{x}_I\overline{R}) \cup (\overline{x}_I\overline{R}, \overline{R})$, investors should not invest in the fund, that is $q_I(r) = 0$. As for $r \in \{\underline{x}_I\underline{R}, \underline{x}_I\overline{R}, \overline{x}_I\underline{R}, \overline{x}_I\overline{R}\}$, investors are indifferent to $q_I(r) = q_r \in [0,1]$, because

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

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 $x^* \in [\underline{x}_I, \overline{x}_I]$. Given the above response of investors, we know that the manager prefers $\sup[\underline{x}_I, \overline{x}_I) = \overline{x}_I$ to any $x \in [0, \overline{x}_I)$. This is because,

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

and

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

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

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

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

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

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

$$\begin{aligned}
&(A\overline{x}_I + \beta)[q_I(0)p(0) + q_I(\overline{x}_I\underline{R})p(\underline{R}) + q_I(\overline{x}_I\overline{R})p(\overline{R})] \\
&\quad = (A\overline{x}_I + \beta)[q_I(0)p(0) + p(\underline{R}) + p(\overline{R})] \\
&\quad \geqslant \sup_{x \in (\overline{x}_I, 1]} (Ax' + \beta)[q_I(0)p(0) + q_I(x'\underline{R})p(\underline{R}) + q_I(x'\overline{R})p(\overline{R})] \\
&\quad = (A + \beta)q_I(0)p(0) \\
&\quad \Leftrightarrow (A\overline{x}_I + \beta)[q_I(0)p(0) + p(\underline{R}) + p(\overline{R})] \geqslant (A + \beta)q_I(0)p(0)
\end{aligned}$$

$$\Leftrightarrow q(0) \leqslant \frac{(A\overline{x}_I + \beta)(1 - p(0))}{A(1 - \overline{x}_I)p(0)}.$$

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

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

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

and

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

A.12 Proof of Proposition 9

Consider the subgames with given historical r, for any a, if $r \in (\overline{x}_I(a)\underline{R}, \underline{x}_I(a)\underline{R}) \cup (\underline{x}_I(a)\overline{R}, \overline{x}_I(a)\overline{R})$, investors with a always should invest in the fund, that is $m(r, a) = 1$. Similarly, given $r \in [\underline{R}, \overline{x}_I(a)\underline{R}) \cup (\underline{x}_I(a)\underline{R}, 0) \cup (0, \underline{x}_I(a)\overline{R}) \cup (\overline{x}_I(a)\overline{R}, \overline{R})$, investors should not invest in the fund, that is $m(r, a) = 0$. As for $r \in \{\underline{x}_I(a)\underline{R}, \underline{x}_I(a)\overline{R}, \overline{x}_I(a)\underline{R}, \overline{x}_I(a)\overline{R}\}$, 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 $\overline{x}_I(a) \in (0, 1)$, $\underline{x}_I(a)$ strictly decreases with a and $\underline{x}_I(a)$ strictly increases with a. It implies that if $m(\overline{R}x, a) = 1$ for some a, then $m(\overline{R}x, a') = 1$ for all $a' \leqslant a$.

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

Then, the expected payoff of the manager is

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

In the second line, we use the fact that the value of $m(\hat{a}, x\overline{R})$ 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_I^*) is that the manager prefers x^* to any $x \in [0, x^*)$

$$\begin{aligned}(x^*A+\beta)&\left[p(0)\int_{\underline{a}}^{\hat{a}(x^*)}1\mathrm{d}F(a)+p(\overline{R})\int_{\underline{a}}^{\hat{a}(x^*)}1\mathrm{d}F(a)+p(\underline{R})\int_{\underline{a}}^{\hat{a}(x^*)}1\mathrm{d}F(a)\right]\\&\geqslant(xA+\beta)\left[p(0)\int_{\underline{a}}^{\hat{a}(x^*)}1\mathrm{d}F(a)+p(\overline{R})\int_{\underline{a}}^{\hat{a}(x)}1\mathrm{d}F(a)+p(\underline{R})\int_{\underline{a}}^{\hat{a}(x)}1\mathrm{d}F(a)\right].\end{aligned}$$

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_I^*) is that the manager prefers x^* to any $x \in (x^*, 1]$

$$\begin{aligned}(x^*A+\beta)&\left[p(0)\int_{\underline{a}}^{\hat{a}(x^*)}1\mathrm{d}F(a)+p(\overline{R})\int_{\underline{a}}^{\hat{a}(x^*)}1\mathrm{d}F(a)+p(\underline{R})\int_{\underline{a}}^{\hat{a}(x^*)}1\mathrm{d}F(a)\right]\\&\geqslant(xA+\beta)\left[p(0)\int_{\underline{a}}^{\hat{a}(x^*)}1\mathrm{d}F(a)+p(\overline{R})\int_{\underline{a}}^{\hat{a}(x)}1\mathrm{d}F(a)+p(\underline{R})\int_{\underline{a}}^{\hat{a}(x)}1\mathrm{d}F(a)\right].\end{aligned}$$

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

A.13 Proof of Proposition 10

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

$$\begin{aligned}\overline{q}(x)&=\sup\left\{q\in[0,1]\middle|k_1(x)q-\frac{1}{2}\int_{\underline{a}}^{F^{-1}(q)}a\mathrm{d}F(a)k_2(x)\geqslant0\right\},\\\underline{q}(x)&=\inf\left\{q\in[0,1]\middle|k_1(x)(1-q)-\frac{1}{2}\int_{F^{-1}(q)}^{\overline{a}}a\mathrm{d}F(a)k_2(x)\leqslant0\right\}.\end{aligned}$$

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

$$k_1(x) \times 0 - \frac{1}{2}\int_{\underline{a}}^{F^{-1}(0)} a\mathrm{d}F(a)k_2(x) = 0, \tag{A19}$$

$$k_1(x) \times (1-1) - \frac{1}{2}\int_{F^{-1}(1)}^{\bar{a}} a\mathrm{d}F(a)k_2(x) = 0, \tag{A20}$$

$$\begin{aligned} \frac{\mathrm{d}k_1(x)q - \frac{1}{2}\int_{\underline{a}}^{F^{-1}(q)} a\mathrm{d}F(a)k_2(x)}{\mathrm{d}q}\Big|_{q=0} &= k_1(x) - \frac{1}{2}F^{-1}(q)k_2(x)\Big|_{q=0} \\ &= k_1(x) - \frac{1}{2}\underline{a}k_2(x) > 0, \end{aligned} \tag{A21}$$

$$\begin{aligned} \frac{\mathrm{d}k_1(x)(1-q) - \frac{1}{2}\int_{F^{-1}(q)}^{\bar{a}} a\mathrm{d}F(a)k_2(x)}{\mathrm{d}q}\Big|_{q=0} &= -\left[k_1(x) - \frac{1}{2}F^{-1}(q)k_2(x)\right]\Big|_{q=0} \\ &= -\left[k_1(x) - \frac{1}{2}\underline{a}k_2(x)\right] < 0, \end{aligned} \tag{A22}$$

$$\frac{\mathrm{d}^2k_1(x)q - \frac{1}{2}\int_{\underline{a}}^{F^{-1}(q)} a\mathrm{d}F(a)k_2(x)}{\mathrm{d}q^2} = -\frac{1}{2f(q)} < 0, \tag{A23}$$

$$\frac{\mathrm{d}^2k_1(x)(1-q) - \frac{1}{2}\int_{F^{-1}(q)}^{\bar{a}} a\mathrm{d}F(a)k_2(x)}{\mathrm{d}q^2} = \frac{1}{2f(q)}k_2(x) > 0. \tag{A24}$$

Combining $k_1(x) - \frac{1}{2}\int_{\underline{a}}^{F^{-1}(1)} a\mathrm{d}F(a)k_2(x) = k_1(x) - \frac{1}{2}\int_{\underline{a}}^{\bar{a}} a\mathrm{d}F(a)k_2(x) < 0$, (A19), (A21) and (A23), we have $\bar{q}(x) \in (0,1)$. Combining $k_1(x)(1-0) - \frac{1}{2}\int_{F^{-1}(0)}^{\bar{a}} a\mathrm{d}F(a)k_2(x) = k_1(x) - \frac{1}{2}\int_{\underline{a}}^{\bar{a}} a\mathrm{d}F(a)k_2(x) < 0$, (A20), (A22) and (A24), we have $\underline{q}(x) = 0$.

Similarly, given $k_1(x) - \frac{1}{2}\int_{\underline{a}}^{\bar{a}} a\mathrm{d}F(a)k_2(x) > 0$, by (A19) - (A24), we have $\underline{q}(x) \in (0,1)$ and $\bar{q}(x) = 1$.

B Institutional Background

This appendix provides examples of intermediaries in delegated asset management in practice, especially featuring personalized advisories and/or recommendation signals.

Fintech platforms, as somehow new entrants in the asset management business, are the closest corresponding to our baseline settings. They highlight their goals of maximizing and protecting users' investment affairs. They put more effort in increasing their technologies and designs for 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.[35] It connects retail investors with alternative investments managed by different fund managers and makes personalized rec-

[35] According to https://www.yieldstreet.com/about/ Date of visit: Sep 09, 2024.

ommendations based on user profiles and investment goals. In particular, as Figure B1 shows, it works in three main steps. First, they *explore* investment opportunities, highlighting their advantages in knowing more opportunities (even including real estate, art, legal finance) and providing easy access. Second, they *invest* with confidence, especially according to users' risk tolerance and the projects' past performance. Third, they track the *earnings* and especially visualize asset allocations.

Commercial banks have a long tradition of providing asset management services. Some businesses have been split into specialized companies or platforms. In this era of digital finance, easy access and low cost in on-line / in-app usage, as well as personalized services, are commonly highlighted. For example, Merrill Guided Investing (under Bank of America), as shown in Figure B2 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 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 within the platform, and users receive investment recommendations based on their profiles and risk tolerance, as Figure B3 shows.

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

Ant's ecosystem comprises five major components: online consumption, mobile payment, investment, consumer credit, and healthcare insurance (for a detailed introduction, see Hong et al. (2020)). The all-in-one ecosystem enables it to conduct analysis of investor preferences, as shown in Figure B4. In the app interface, one can see mutual fund recommendation pages, as depicted in Figure B5.

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 different risk assessment results. This includes determining the investment objectives, how to allocate assets, and the expected investment returns. Following the risk assessment and investment goals, BangNiTou works by recommending a portfolio selected from 6,000 mutual funds (see Figure B6). In 2021, assets under BangNiTou sits at ¥6.9 billion (about $1 billion) (Bloomberg, 2021).

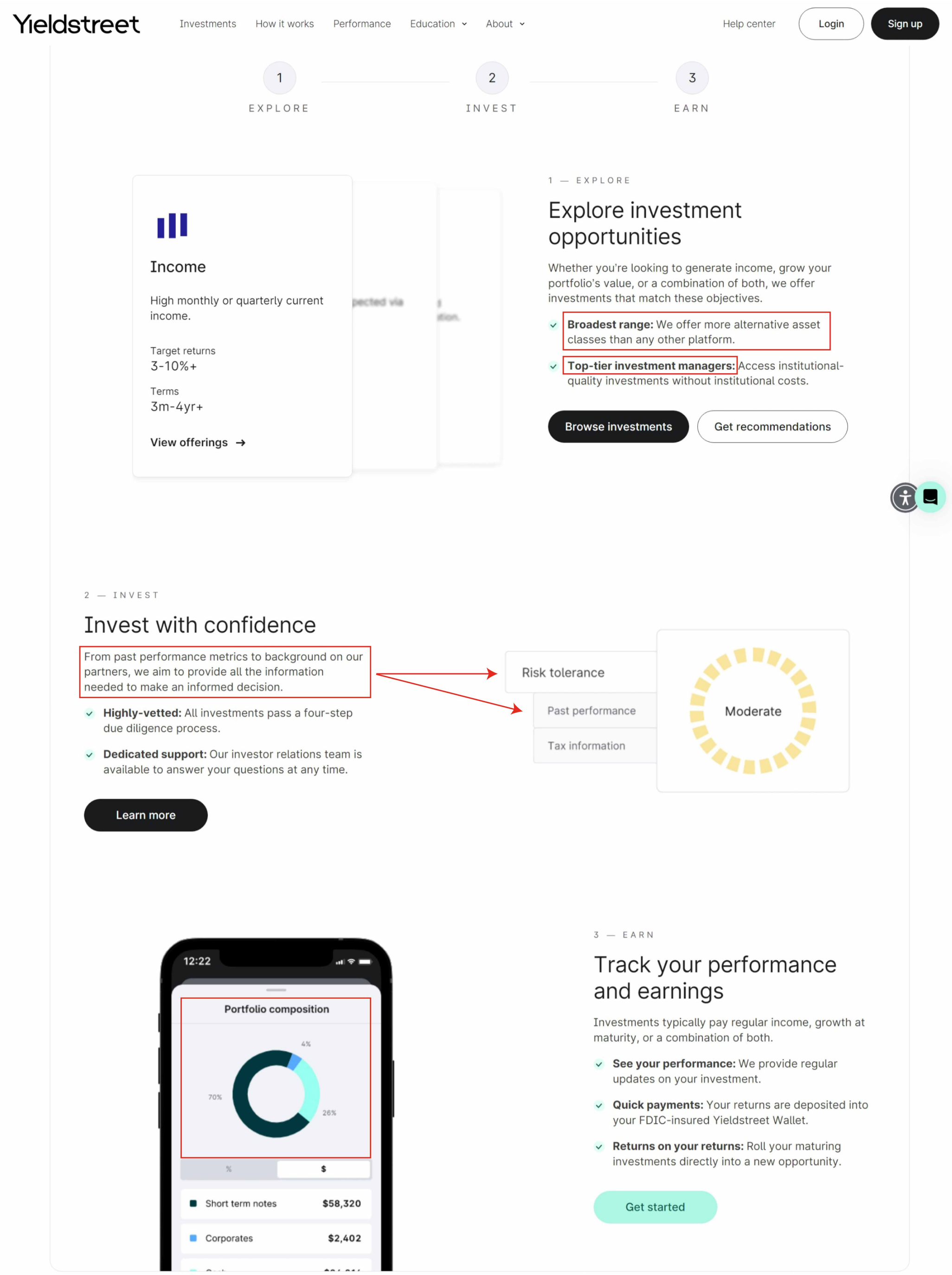


Figure B1: Delegated Investment service of Yieldstreet

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

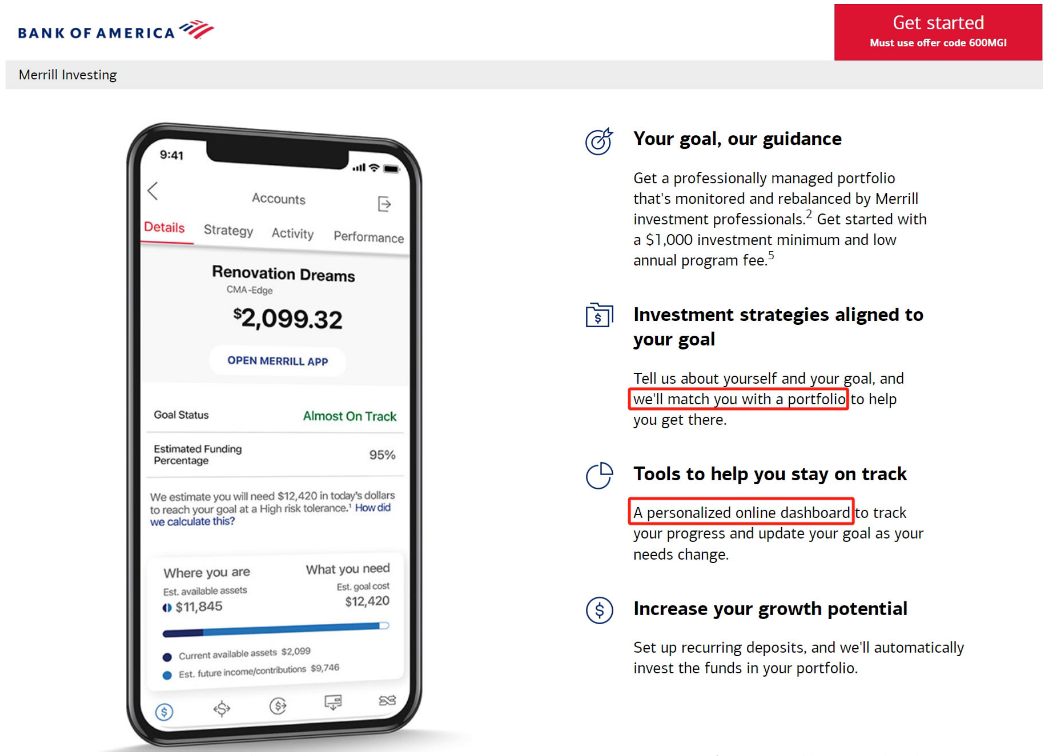


Figure B2: Personalized investment matching service of Merrill Guided Investing

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

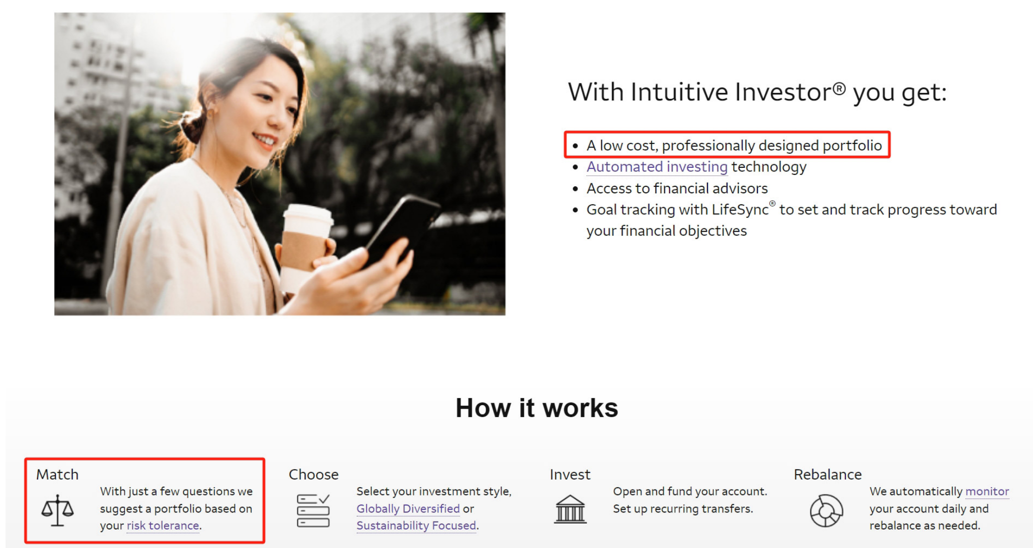


Figure B3: Intuitive Investor's Personalized investment based on risk tolerance

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

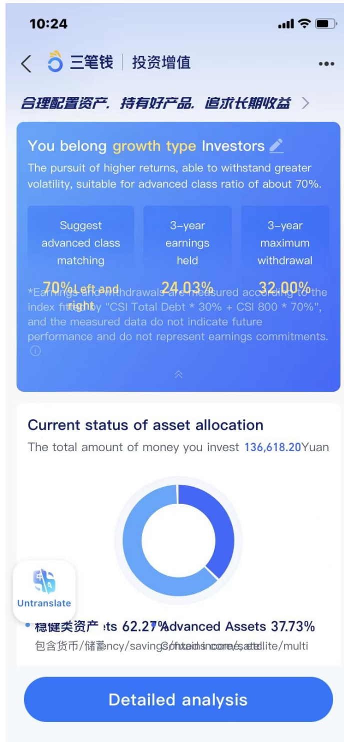
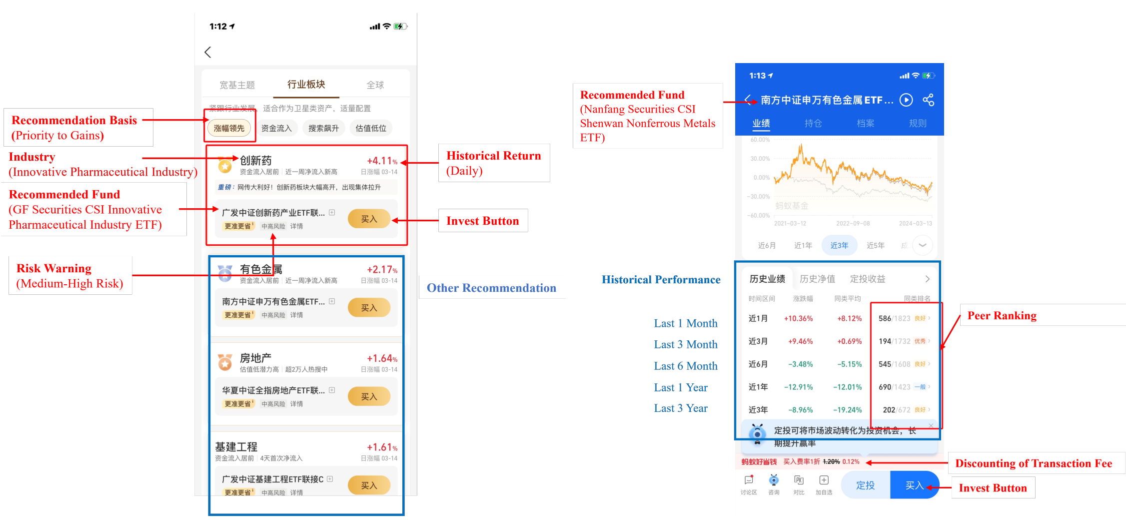
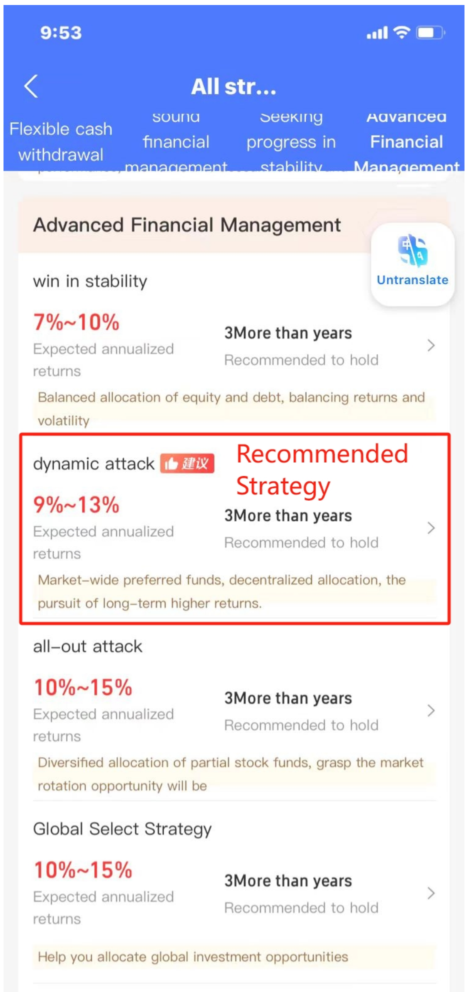


Figure B4: Analysis of Risk Tolerance and Investment Object in Ant Financial

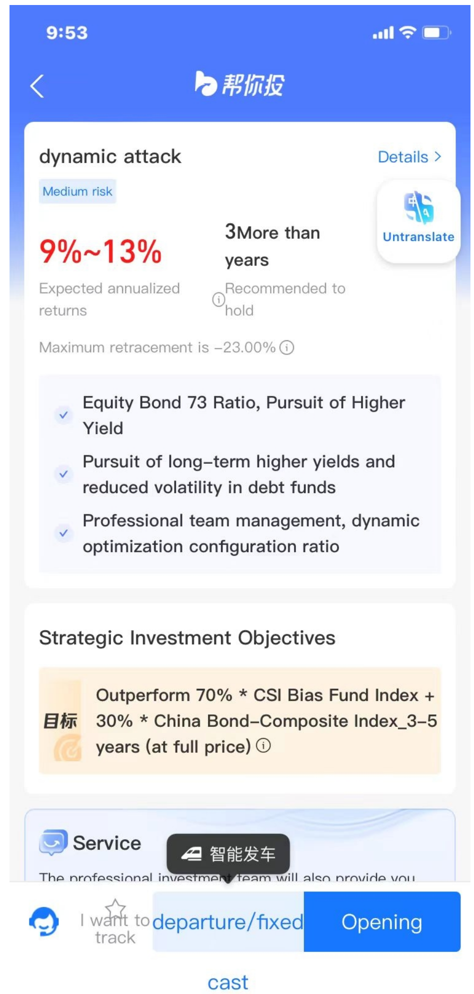


(a) Recommended Funds **(b)** Detail of A Recommended Fund

Figure B5: Recommendation service of Ant Group Co. Ltd.



(a) Recommended Strategy



(b) Detail of the Strategy

Figure B6: Robo-advisory service of BangNiTou