

Searching for and evaluating outsourced chief investment officers

Abstract: This study analyzes the impact of the outsourced chief investment officer (OCIO) index and OCIO search consultant on asset owners' selection of OCIOs. We adopt an agent-based model with a reinforcement learning method by defining the OCIO search problem as a multi-armed bandit problem. Our model highlights the significance of the information regarding potential managers provided by the OCIO index and an OCIO search consultant. Our simulation results indicate that the more managers in the market, the harder it is for asset owners to identify the optimal OCIO. If asset owners place greater emphasis on their preferences when evaluating OCIOs, those challenges might be eased. Our results suggest that OCIO search consultants and the OCIO index can support asset owners' decision-making.

Keywords: Agent-based simulation; Outsourced chief investment officer (OCIO); OCIO index; Multi-armed bandit problem; Search consultant

JEL Classification: C61; D83; G29

1. Introduction

Outsourced chief investment officer (OCIO) services are a type of delegated asset management. However, unlike traditional asset management approaches, OCIOs receive authority over strategic asset management from asset owners (Clark and Urwin, 2019). Asset allocation can be categorized into two types—namely, strategic and tactical asset allocation (Anson, 2004). Strategic asset allocation involves determining the long-term composition of assets and the range of variation in the medium term to achieve financial goals. By contrast, tactical asset allocation involves executing changes in the portfolio—as determined by strategic asset allocation in the medium term—based on the investment outlook. Traditionally, delegated asset management involves delegating authority only to managers for tactical asset allocation. However, OCIOs typically delegate broad authority over strategic asset allocation. Therefore, OCIOs provide comprehensive asset management services with a wider scope of delegation than existing asset management services. Considering the increasing uncertainty in the global financial market, asset owners who are not expert managers have difficulty managing their assets. The asset under management of the global OCIO industry increased from USD 1.29 trillion in 2016 to USD 2.46 trillion in 2021. This figure is expected to increase to greater than USD 3 trillion by 2023.¹ As the OCIO market grows, matching asset owners with appropriate OCIO providers has become increasingly important. However, the presence of numerous OCIO providers makes it challenging for asset owners to identify OCIOs suitable for their asset management objectives.

¹ See “The Growing OCIO Industry Intensifies Market Competition for Managers and Consultants” - January 25, 2022 – Institutional Investor.

Conducting elaborate search and evaluation are important for selecting the optimal OCIO. The ideal approach to identifying the optimal OCIO is for the asset owner to consider all available managers. However, this approach has practical limitations in real-world markets. Asset owners must invest resources each time they contact a manager, and comparing alternatives and selecting them requires time and money. Information asymmetry between managers and asset owners is an issue. Managers have expert portfolio management knowledge, whereas asset owners have less expertise. As such, evaluating a manager's performance perfectly is difficult for an asset owner. An owner can observe the realized return, but determining whether the portfolio risk has been managed appropriately is more challenging. If this evaluation is incorrect, identifying the optimal OCIO becomes more difficult. Information asymmetry also reduces transparency in the OCIO industry and can cause agency problems.

Several evaluation methods have been proposed for increasing transparency in the OCIO industry (Park and Ryu, 2023). We consider two of these methods—specifically, the OCIO index and OCIO search consultants. Nasdaq created the Alpha Nasdaq OCIO Index with Alpha Capital Management. This index presents the average performance of OCIO providers and helps asset owners objectively measure OCIOs' performance through direct comparisons. OCIO search consultants help asset owners select and evaluate OCIOs' performance. The consultants identify the characteristics of their clients (i.e., asset owners) and search for the optimal OCIOs. Unlike clients, OCIO search consultants specialize in seeking and evaluating OCIOs, suggesting that they are relatively free from the constraints faced by clients. Moreover, they have expertise in evaluating OCIOs. For example, North Pier—an OCIO search consultant—created the Discretionary Investment Management Working Group on Data Standards to provide customers with accurate and detailed comparisons of OCIOs. An OCIO search consultant plays a role like that of a real estate broker by interconnecting asset owners with OCIO providers and assisting them in their searches.

We analyze the impact of these two evaluation methods on asset owners' OCIO choices. Specifically, we construct an agent-based model using a reinforcement learning approach. We define each asset owner's search for the optimal OCIO as a multi-armed bandit problem as follows: Asset owners face the challenging task of identifying the most suitable OCIO providers to meet their asset management goals. One approach to finding the right OCIO is through trial and error using a process called reinforcement learning, which involves an agent (in this case, the asset owner) who decides whether to continue with an option that provided high rewards in the past (exploitation) or explore a new alternative. During exploration, the agent randomly selects an alternative, meaning no prior information exists regarding the managers. To overcome this information gap, asset owners can use the OCIO index and search for consultants to obtain information regarding managers. During exploration, asset owners who refer to the OCIO index choose managers randomly from among those with high investment performance. By contrast, asset owners who hire an OCIO search consultant receive a list of recommended managers from the consultant. In summary, using the OCIO index and search

consultants, asset owners can increase their chances of identifying the right OCIO provider without relying on a trial-and-error approach.

Through simulation analyses, we obtain the following key results. The likelihood of asset owners identifying the optimal OCIO decreases as the number of managers increases, especially compared to the effect of decreasing weight. This trend poses challenges to selecting the right OCIO in a competitive market. Considering personal preferences during the selection process can help overcome these challenges. Additionally, using OCIO search consultants and the OCIO index can offer useful insights to asset owners in determining the optimal provider.

The remainder of this paper is organized as follows: Section 2 reviews the related literature. Section 3 describes the simulation model. Section 4 presents the results. Section 5 concludes the paper.

2. Literature Review

We use several methodologies to analyze an asset owner's search for the optimal OCIO. The first is the optimal stopping theory. The secretary problem, which describes the optimal stopping theory, is like an asset owner's search for an OCIO. The secretary problem's purpose is to determine a stopping rule that selects the optimal agent from a series of rankable agents. Generally, the stopping rule is derived to maximize the expected utility from limited information (Ferguson, 1989; Palley and Kremer, 2014). Considering that obtaining prior information regarding OCIO funds is difficult, conducting a search directly or visiting an OCIO provider for consultation is necessary. However, information on all OCIOs cannot be gathered simultaneously; thus, the fund prioritizes its search according to a certain standard and selects OCIOs to maximize its expected utility. This pattern is like that observed in secretary problems. Additionally, this analysis is related to Weber and Zheng's (2007) model of a search intermediary. Specifically, they use a search theory model to analyze situations wherein individual consumers seek products to purchase through search intermediaries, such as Google or Yahoo. The consumer compares the search cost with the marginal expected utility from conducting another search and determines the number of products to be considered based on the rankings of products posted by the search broker.

A problem that naturally arises in this setting is the expected utility's realistic estimation. The problem can be resolved by adopting reinforcement learning. Thus, an asset owner's search for the optimal OCIO is similar to a multi-armed bandit problem, which involves identifying the option that provides the highest expected reward among several options (Sutton and Barto, 2018). This problem can be illustrated using a slot machine with multiple handles—each with a different probability distribution for winning. The knobs' expected returns can be estimated by pulling them and measuring the rewards. The multi-armed bandit problem is a simple yet powerful framework for addressing sequential stochastic decision problems (Marković, Stojić, Schwöbel, and Kiebel, 2021). Several studies extend and apply this idea to numerous problems. Some studies propose improving the method

of calculating the reward based only on the expected return to calculate the reward through a risk measure, which is useful in portfolio selection problems (Khajonchotpanya, Xue, and Rujeerapaiboon, 2021). Other studies extend the multi-armed bandit problem to a multi-agent problem wherein the agents' cooperative behavior is considered (Landgren, Srivastava, and Leonard, 2021; Madhushani and Leonard, 2020).

We refer to this reinforcement learning method as one by which the asset owner finds the optimal OCIO. We assume that asset owners use two criteria when selecting an OCIO—specifically, the OCIO's investment performance and their preferences for OCIOs. Performance refers to the rate of return during the OCIO's contract period, whereas preferences cover all other factors. The OCIO's rate of return is the sum of the expected returns and random shocks. Asset owners calculate the rewards from an OCIO based on performance and preferences and, thereafter, select an OCIO to maximize the expected reward.

We assume that asset owners estimate the average reward from an OCIO as the sample mean reward over the past period and select the OCIO with the greatest sample mean. It is known as the sample average approximation method. An agent that uses this method determines the optimal action through exploitation and exploration. Exploitation—also called greedy action—means that the agent selects the action with the greatest sample mean. However, an agent who behaves greedily continues selecting only the first action. To find the optimal action, the greedy action must be abandoned and other actions must be selected. Relinquishing a greedy action with a certain probability and randomly choosing another action is called the ϵ -greedy method.

We analyze the impacts of these two methods on asset owners' OCIO selection by constructing an agent-based model. An agent is given the ability to interact and autonomously judge and makes decisions in a virtual space (Axtell, Axelrod, Epstein, and Cohen, 1996). An agent-based model can easily incorporate heterogeneous agents (Rahmandad and Sterman, 2008). With limited information, observing other agents' decisions is difficult for each agent (Park, Hong, and Ryu, 2023). In turn, deriving equilibria mathematically is challenging because the interactions among agents are complex (Epstein, 1999). Agent-based simulation involves constructing a complex agent-based model and observing patterns in the agents' behavior through simulation (Cristelli, 2013).

In this study, we modify the ϵ -greedy method to compare the impact of the OCIO index and the OCIO search consultant. Using the standard ϵ -greedy method, agents select an action randomly when they explore, implying that they possess no information regarding the problem. We assume that the OCIO index and search consultants provide asset owners with additional information on OCIOs. The OCIO index provides agents with the average returns of all OCIOs, while an OCIO search consultant ranks managers based on the rewards that they receive.

3. Model

We construct a model based on the multi-armed bandit problem. Learning within the multi-armed bandit problem is considered the simplest version of reinforcement learning. Reinforcement learning typically requires three elements—specifically, action, reward, and state (Park and Ryu, 2022). When an agent selects an action, a reward is provided and the subsequent state is determined. The agent selects an action that maximizes the expected reward in a given state. In reinforcement learning, the optimal action differs according to the state. However, the multi-armed bandit problem is a reinforcement learning problem with only one state, suggesting that the agent has only one optimal action. We define the asset owner’s OCIO search problem as a multi-armed bandit problem. These problems are similar in several ways: They offer several options, each option provides the agent with a random reward, and the agent estimates the options’ true rewards by repeatedly selecting them. Table 1 summarizes the notations used in this study.

[Table 1 about here]

The asset owner action set is a list of managers. When the asset owner selects a manager as the OCIO, the owner receives a reward from the OCIO. As the primary purpose of asset owners hiring OCIOs is to provide competent asset management services, investment performance is an important component of the reward. However, the evaluation criteria may include factors other than investment performance. Asset owners ask managers about their investment philosophies.² For example, some asset owners may ask the OCIO to include companies whose management strategies are based on environmental, social, and governance (ESG) principles. Other asset owners may wish to include managers who adhere to diversity, equity, and inclusion (DEI) principles in the portfolio. These requirements can be considered preferences that are unrelated to investment performance. We define a reward when asset owner i chooses manager a at time t ($R_{i,t}(a)$) as a linear combination of preference and performance as in Equation (1):

$$R_{i,t}(a) = w \cdot u_{a,i} + (1 - w) \cdot r_{a,t}, \quad (1)$$

where $u_{a,i}$ denotes the asset owner i ’s preference for manager a . $r_{a,t}$ denotes the return of manager a ’s portfolio at the t -th iteration. w is the weight on preference.

Referring to Hotelling (1929)’s model, we calculate the asset owner’s preference for an OCIO based on the distance between their preference parameters, as indicated in Equation (2). An asset owner’s preference parameter refers to the degree of requirement for an investment philosophy, such as

² See “OCIO Managers Expand ESG, DEI Prowess to Meet Requests” - January 04, 2022 – Pension & Investments

ESG and DEI, while that of an OCIO refers to the degree of adherence to such principles:

$$u_{a,i} = -|\tau_a - \psi_i|, \quad (2)$$

where τ_a and ψ_i denote preference parameters of manager a and asset owner i , respectively.

We define an OCIO's return based on traditional portfolio theory. Each OCIO finds a tangent portfolio based on all available assets. They derive a capital allocation line based on their tangent portfolio and risk-free returns and select a portfolio that maximizes the asset owner's utility. Asset owners prefer portfolios with higher expected returns than variance (risk). Considering that a relative comparison between managers is important, the variance has been normalized to one. The OCIO's rate of return consists of the expected return and variance, as indicated in Equation (3):

$$r_{a,t} = \mu_a + \tilde{z}_t, \quad \tilde{z}_t \sim N(0,1), \quad (3)$$

where μ_a and \tilde{z}_t denotes the expected return of manager a 's portfolio and random shock at the t -th iteration, respectively. \tilde{z}_t follows a standard normal distribution.

The optimal OCIO for agent i is the manager who provides the highest expected reward. The expected reward is calculated using Equation (4). Considering that the variance follows a standard normal distribution, the expected reward is the weighted sum of the asset owner's preference and the expected return on the manager's portfolio:

$$E[R_{i,t}(a)] = w \cdot u_{a,i} + (1 - w) \cdot \mu_a. \quad (4)$$

Each agent estimates the expected reward using the sample-average approximation method. The estimate for the expected reward $Q_{i,t}(a)$ is defined as in Equation (5), where $A_{i,t}$ means the agent i 's action at the t -th iteration and $I(\cdot)$ is the indicator function. Asset owner i chooses action $A_{i,t}$ that maximizes $Q_{i,t}$:

$$Q_{i,t}(a) = \frac{\sum_{k=1}^{t-1} R_{i,k}(a) \cdot I(A_{i,k}=a)}{\sum_{k=1}^{t-1} I(A_{i,k}=a)}. \quad (5)$$

We adopt the ϵ -greedy method. Agents select the greedy action with probability $1 - \epsilon$ and explore with probability ϵ . We adopt a decaying ϵ , as in Equation (6), to ensure that the asset owner actively explores in the beginning and acts greedily after a certain amount of learning:

$$\epsilon_t = \exp\left(-\frac{T}{10} \cdot t\right), \text{ where } t = 1, \dots, T. \quad (6)$$

We modify the ϵ -greedy method to compare situations wherein asset owners can hold different amounts of information. In the standard scenario, agents randomly select actions, implying that they have no information to rank the possible actions. However, if agents possess preliminary information, they can explore more efficiently. We assume that the OCIO index and search consultants offer agents additional information in different ways. The OCIO index provides OCIOs' average returns at time $t - 1$ and agents can randomly select from the OCIOs who outperformed the OCIO index during exploration. The OCIO search consultants average the historical rewards from each OCIO. Referring to Yinger's (1981) real estate broker model, we assume that the OCIO search consultants provide a list of managers in order of the highest rewards. Asset owners can randomly select an asset from the list.

4. Simulation Results

4.1. Simulation settings

We conduct simulations using the model described in Section 3. The simulation settings for this paper are as follows³: Each simulation involves a market with m managers and n asset owners, with the number of asset owners (n) fixed at 100. However, we vary the number of managers (m), depending on the scenario. In the baseline scenario, there are 10 managers. When there are fewer managers in the market, we reduce the number to five, which is half of the baseline value. Conversely, in scenarios with several managers, we increase the number to 20, which is double the baseline value. At the beginning of each simulation, the parameters are randomly assigned to asset owners and managers. Asset owners are assigned a preference parameter (ψ_i) uniformly distributed between 0 and 1. Managers are randomly assigned a preference parameter (τ_a) and an expected return (μ_a). Additionally, these two parameters are distributed uniformly between 0 and 1. Once the agents' parameters are determined, the manager with the highest expected reward is determined as the optimal OCIO for an asset owner, according to Equation (4). The weight assigned to the preference and performance when calculating the reward is an important parameter. In the baseline scenario, the weight of preference (w) is 0.5. We vary the weight to 0.25 and 0.75 in the other scenarios. The former implies that asset owners value performance more, whereas the latter implies that they value preferences more.

After setting the parameters, the asset owners select and evaluate the OCIOs at each iteration using the following process. At the t -th iteration's beginning, the asset owners select an OCIO based on the Q -value for each manager accumulated in the previous time. However, with a probability of $1 - \epsilon$,

³ In the process of summarizing the simulation algorithms, we have significantly referenced the literature introduced by Huang, Ma, and Fu (2018), including Huang et al. (2018), Kan et al. (2018), Li et al. (2018a), Li et al. (2018b), Meng et al. (2018), and Vamvakas, Tsiropoulou, and Papavassiliou (2018).

they explore some of the managers randomly. We use three exploration methods here. The first is self-directed searching wherein the asset owner randomly selects one among all managers—similar to a simple ϵ -greedy algorithm—with no additional information regarding the optimal OCIO. This method serves as a benchmark for the following two exploration methods: In the second method, asset owners use the OCIO index. The OCIO index is the average of the performance of OCIOs from one previous period. Asset owners select the manager who outperforms the OCIO index as the subsequent OCIO. The third method involves searching for OCIOs using an OCIO search consultant. In this method, OCIO search consultants average the past performance of each manager. They compile a list of managers based on their average past performance and recommend some from the list as candidates for the optimal OCIO. Asset owners randomly select one recommended manager. A higher percentage of asset owners successfully identifying their best OCIO indicates better performance of the exploration method.

Managers’ portfolio performances are disclosed after the OCIO selection process. Asset owners observe the OCIO’s performance in calculating the reward for their OCIO choice based on Equation (1). Based on these rewards, asset owners update their Q -value. Each simulation runs for 1,000-time steps ($T = 1,000$). Thereafter, the simulation is repeated 100 times for each scenario to ensure statistical significance. We use a specific random seed number for each simulation to make fair comparisons among the three exploration methods. We implement the simulations using Matlab and control the random number generation with the ‘*rng*’ function.

We conduct simulations for five scenarios by varying the weights of the preferences and the numbers of managers. The baseline scenario has equal weights for preference and performance ($w = 0.5$), and 10 managers ($m = 10$). In subsection 4.2, we compare the baseline scenario with two scenarios that differ in the number of managers ($m = 5$ or $m = 20$) with the baseline weight on preference ($w = 0.5$). In subsection 4.3, we compare the baseline scenario with two scenarios that differ in the weight of preference ($w = 0.25$ or $w = 0.75$), with the baseline number of managers ($m = 10$).

4.2. Number of managers

This section analyzes the impact of the changes in the number of managers on the percentage of asset owners who identify the optimal OCIOs. We use the baseline scenario of 10 managers and compare the simulation results with scenarios that have half as many managers and twice as many managers. The weight for preference is fixed at 0.5 in the three scenarios. The simulation results are presented in Figure 1, which depicts the percentage of asset owners who successfully identify their optimal OCIOs as the number of managers changes.

[Figure 1 about here]

Figure 1 shows that all exploration methods reach a certain level of performance convergence,

irrespective of the number of managers involved. During the early stages of the learning process, compared to the self-directed search and OCIO search consultant methods, the OCIO index method provides the worst results. At around the 300th iteration, the self-directed search method catches up with the OCIO index method, and at around the 400th iteration, both methods catch up with the OCIO search consultant method. This implies that the OCIO index method converges more slowly than the other two methods. In contrast, the OCIO search consultant method outperforms the self-directed search method, regardless of the number of managers in all periods.

For all exploration methods, the optimal selection proportion seems to decrease as the number of managers increases. However, as the number of managers increases, the performance of the models with additional information is relatively better than that of the self-directed search method. To observe the impact of the number of managers on the performance of the exploration method more intuitively, Figure 2 summarizes the results by dividing them according to each exploration method.

[Figure 2 about here]

Figure 2 reveals that the percentage of asset owners finding the optimal OCIO decreases for all periods as the number of managers increases, irrespective of the exploration method used. However, the amount of the reduction varies by the method used. When the number of managers increases from 10 to 20, the OCIO index exhibits a more significant decrease than when the number of managers increases from 5 to 10. Conversely, the OCIO search and consultation method demonstrates a relatively consistent decline. Seemingly, the results of the self-directed search method deteriorated more significantly than those of the other two methods. Like the OCIO index method, seemingly, a more substantial decrease occurs when the number of managers increases from 10 to 20 than when it increases from 5 to 10. To compare the methods directly, we compute the difference between the optimal selection proportions obtained using the two methods. Figure 3 illustrates the difference in the proportions between the two methods.

[Figure 3 about here]

Figure 3 enables us to thoroughly investigate the findings depicted in Figures 1 and 2. The OCIO index method initially exhibits lower performance than the self-directed searching method but exhibits better performance at a certain point. In particular, the degree of performance improvement significantly increases as the number of managers increases in the latter stages of the learning process. As presented in Figure 1, the OCIO search consultant method exhibits better overall performance than the self-directed searching method. Additionally, we observe that the performance gap between the two methods increases with the number of managers. However, unlike the OCIO index method, the

difference in performance increases more significantly when the number of managers increases from 10 to 20 than when it increases from 5 to 10. Owing to these differences, the comparison between the OCIO index and OCIO search consultant methods reveals that the gap between the two is generally the largest when the number of managers is 10.

4.3. Weight on preference

We compare the simulation results based on changes in weight with respect to the preferences. In the baseline scenario, the weight of preference is 0.5. Therefore, asset owners perceive preference and performance as equally important. In the performance scenario, the weight of the preference is 0.25. In this scenario, asset owners assign greater weight to performance. In a preference scenario, wherein the weight of preference is 0.75, asset owners believe that preference is more important than performance. Figure 4 displays the percentage of asset owners who identify their optimal OCIOs as the weight of preference changes.

[Figure 4 about here]

The results in Figure 4 are considerably similar to those in Figure 1, with each exploration method converging well. The OCIO index method converges more slowly than the other two methods, and the OCIO search consultant method dominates self-directed searching in all periods. Although the impact is small compared to the number of managers, the weight of preference consistently affects the percentage of the optimal OCIO selection. Increasing the weight for preference generally increases the results' accuracy. However, unlike changes in the number of managers, weight changes seemingly affect result volatility. Decreasing the weight for performance can reduce the volatility of the results because it decreases randomness in the reward calculation. To gain a more intuitive understanding of the impact of preference weight, we summarize the results according to each exploration method in Figure 5.

[Figure 5 about here]

In Figure 5, increasing the weight of the preference significantly improves accuracy in all methods. However, unlike the changes in the number of managers, weight changes do not lead to method-specific differences. They generally exhibit similar levels of change across methods. To observe these results rigorously, Figure 6 illustrates the difference in the proportions between the two methods. Figure 6 reveals that the confidence intervals of each line overlap, which differs from Figure 3. These results suggest that the weight changes generally do not lead to significant differences between exploration methods.

[Figure 6 about here]

To summarize the simulation results, as the number of managers increases, the percentage of asset owners identifying the optimal OCIO decreases, particularly as the weight of preference decreases. The impact of increasing the number of managers on the decrease in optimal selection is more significant than that of decreasing the weight. Furthermore, an increase in the number of managers has a more pronounced effect on the gap between exploration methods, whereas a change in weight does not.

Our findings have several important implications for future studies. As the OCIO market expands and the number of providers increases, asset owners may face greater challenges in identifying the best OCIO to meet their needs. However, if the OCIO selection process places greater emphasis on factors beyond performance, such as personal preferences, it may mitigate the impact of an increasing number of providers on the difficulty of finding the right OCIO. In this process, the additional information that asset owners can obtain from OCIO search consultants and the OCIO index will significantly contribute to helping asset owners identify the optimal OCIO.

5. Conclusion

This study analyzes the impact of the OCIO index and search consultants on asset owners' OCIO selection using an agent-based model with a reinforcement learning algorithm. We define the OCIO search problem as a multi-armed bandit problem and use prior information on managers from the OCIO index and search consultants by modifying the ϵ -greedy method. Our results reveal that as the number of managers increases, the percentage of asset owners identifying the optimal OCIO decreases. The impact on the optimal selection is more significant as the number of managers increases, compared to decreasing the weight. This poses challenges for asset owners in identifying the right OCIO in a growing market. Incorporating personal preferences into the selection process can mitigate this difficulty. Additionally, using OCIO search consultants and the OCIO index can provide valuable information to asset owners in their search for an optimal provider.

Thus, this study demonstrates the usefulness of the OCIO index and search consultants as viable alternatives to select the optimal OCIO in a competitive market. However, this model can be improved in several ways to provide a comprehensive understanding of the OCIO market. First, asset owners should be allowed to adaptively select their actions, particularly when selecting the exploration method. By analyzing the method that is more likely to be selected, we can gain deeper insights into the selection process. Second, we recommend introducing risk aversion into the model, because the current model assumes that managers' portfolios exhibit the same volatility. Considering that risk management is a significant reason to select an OCIO, further research is required to compare

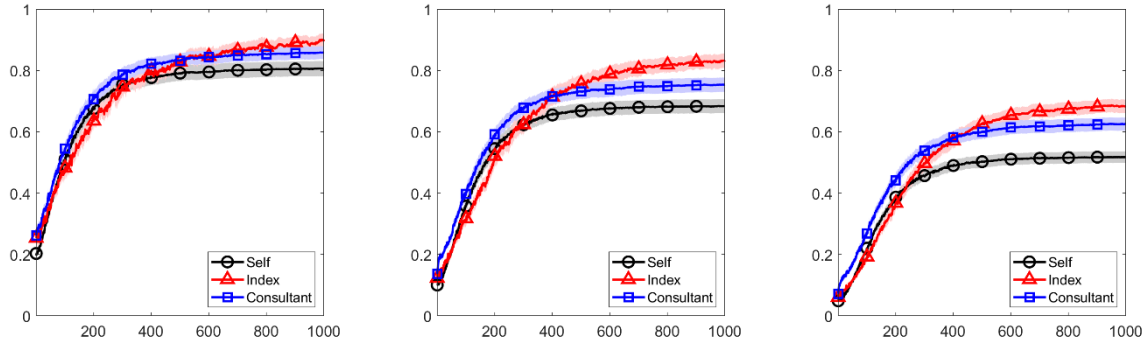
alternatives with varying volatilities.

References

- Anson, M.J. (2004). Strategic versus tactical asset allocation. *Journal of Portfolio Management*, 30(2), 8-22.
- Axtell, R., R. Axelrod, J.M. Epstein, and M.D. Cohen, (1996). Aligning simulation models: A case study and results. *Computational and Mathematical Organization Theory*, 1(2), 123-141.
- Charpentier, A., Elie, R., and Remlinger, C. (2021). Reinforcement learning in economics and finance. *Computational Economics*, 1-38.
- Clark, G.L., and R. Urwin, (2017). The outsourced chief investment officer model of management and the principal-agent problem. *Journal of Retirement*, 4(3), 28-41.
- Cristelli, M. (2013). *Complexity in Financial Markets: Modeling Psychological Behavior in Agent-based Models and Order Book Models*. Springer Science and Business Media.
- Epstein, J.M. (1999). Agent-based computational models and generative social science. *Complexity*, 4(5), pp. 41-60.
- Ferguson, T.S. (1989). Who solved the secretary problem? *Statistical Science*, 4(3), 282-289.
- Hotelling, H. (1929). Stability in competition. *Economic Journal*, 39(153), 41-57.
- Khajonchotpanya, N., Y. Xue, and N. Rujeerapaiboon, (2021). A revised approach for risk-averse multi-armed bandits under CVaR criterion. *Operations Research Letters*, 49(4), 465-472.
- Huang, X. L., Ma, X., and Hu, F. (2018). Editorial: Machine learning and intelligent communications. *Mobile Networks and Applications*, 23, 68-70.
- Landgren, P., V. Srivastava, and N.E. Leonard, (2021). Distributed cooperative decision making in multi-agent multi-armed bandits. *Automatica*, 125, 109445.
- Madhushani, U., and Leonard, N. E. (2020). A dynamic observation strategy for multi-agent multi-armed bandit problem. In *2020 European Control Conference (ECC), IEEE*, 1677-1682.
- Marković, D., H. Stojić, S. Schwöbel, and S.J. Kiebel, (2021). An empirical evaluation of active inference in multi-armed bandits. *Neural Networks*, 144, 229-246.
- Palley, A.B., and M. Kremer, (2014). Sequential search and learning from rank feedback: Theory and experimental evidence, *Management Science*, 60(10), 2525-2542.
- Park, D., J. Hong, and D. Ryu, (2023). Heterogeneous expectations in the housing market: A Sugarscape agent-based model. *Working Paper*.
- Park, D. and D. Ryu, (2022). Supply chain ethics and transparency: An agent-based model approach with Q-learning agents. *Managerial and Decision Economics*, 43(8), 3331-3337.
- Park, D., and D. Ryu, (2023). Small-sized asset owners' OCIO selections and evaluations. *Journal of Derivatives and Quantitative Studies*, Forthcoming.
- Rahmandad, H., and J. Sterman, (2008). Heterogeneity and network structure in the dynamics of

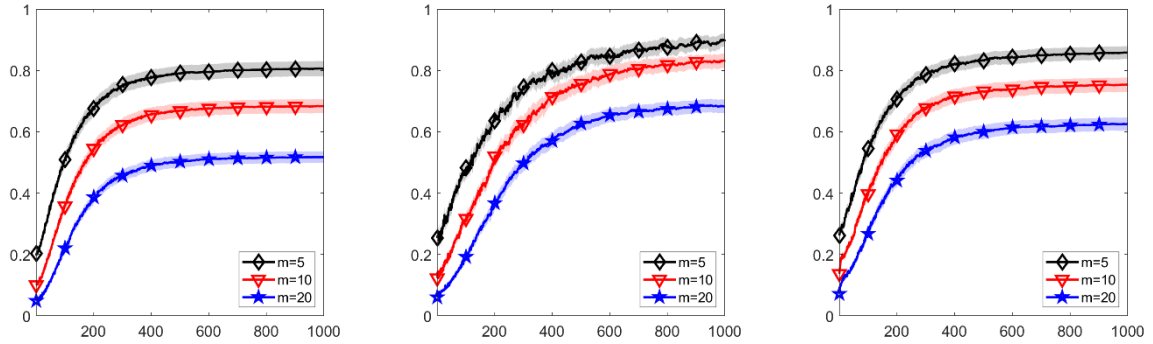
- diffusion: Comparing agent-based and differential equation models. *Management Science*, 54(5), 98-1014.
- Sutton, R.S., and A.G. Barto, (2018). *Reinforcement Learning: An Introduction*. MIT Press.
- Weber, T.A., and Z. Zheng, (2007), A model of search intermediaries and paid referrals, *Information Systems Research*, 18(4), 414-436
- Yinger, J.A. (1981), Search model of real estate broker behavior, *American Economic Review*, 71(4), 591-605.

Figure 1. Proportion of the optimal OCIO selection based on changes in the number of managers
Panel A. With fewer managers ($m = 5$) *Panel B.* Baseline scenario ($m = 10$) *Panel C.* With many managers ($m = 20$)



Note: This figure presents the proportion of the optimal OCIO selection based on the changes in the number of managers when the weight on preference is fixed at 0.5 ($w = 0.5$). Panel A, B, and C present the results in scenarios with 5, 10, and 20 managers, respectively. The vertical axis represents the proportion of asset owners who select the optimal OCIO, while the horizontal axis represents the iteration number. “Self,” “Index,” and “Consultant” refer to different methods of exploration—specifically, self-directed searching, using the OCIO index, and consulting with an OCIO search consultant, respectively. Each line on the graph represents the average of 100 simulation results for a given scenario. The lines with circle, triangle, and square markers indicate the results for scenarios using the self-directed searching method, OCIO index method, and OCIO search consultant method, respectively. The shading around each line represents a 95% confidence interval.

Figure 2. Comparison of the changes in the number of managers by the exploration method
Panel A. Self-directed *Panel B.* OCIO index *Panel C.* OCIO search consultant



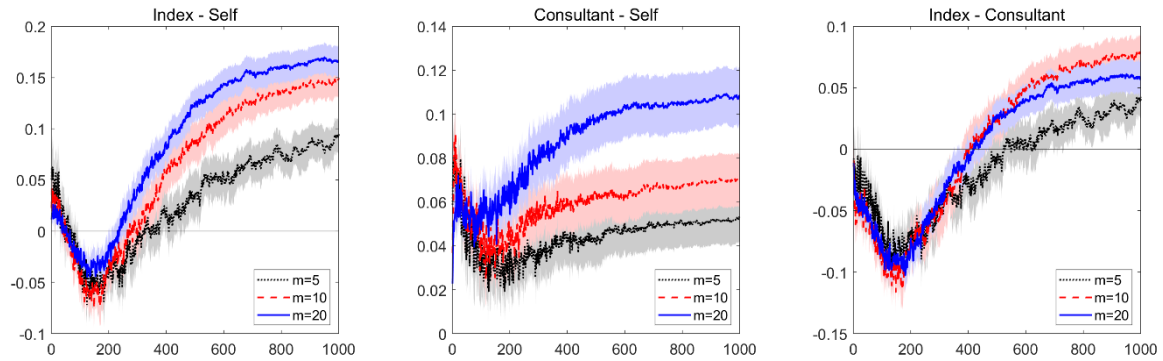
Note: This figure compares the changes in the number of managers by the exploration method when the weight on preference is fixed at 0.5 ($w = 0.5$). Panel A, B, and C present the results using scenarios with self-directed searching, the OCIO index, and the OCIO search consultant, respectively. The vertical axis represents the proportion of asset owners who have chosen the best OCIO, while the horizontal axis represents the iteration number. Each line on the graph represents the average of 100 simulation results for a given scenario. The lines with diamond-shaped, inverted-triangle, and star-shaped markers indicate the results for scenarios with 5, 10, and 20 managers, respectively. The shading around each line represents a 95% confidence interval.

Figure 3. Differences in the optimal OCIO selection proportions between exploration methods based on changes in the number of managers

Panel A. OCIO index and self-directed searching

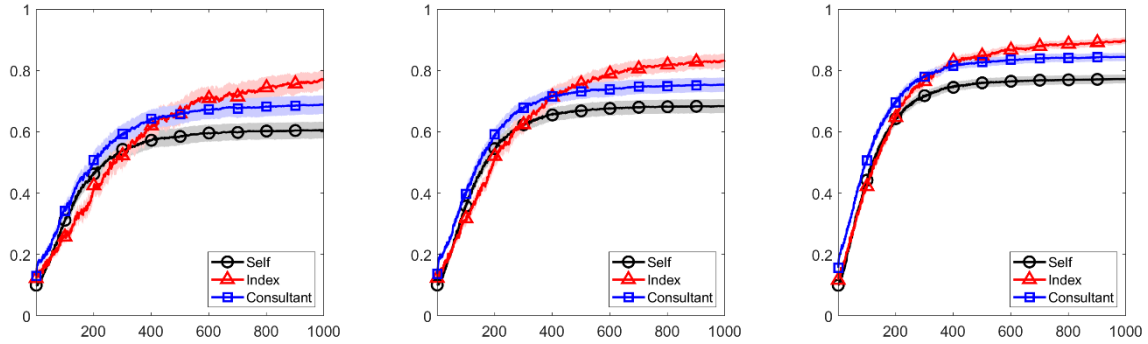
Panel B. OCIO search consultant and self-directed searching

Panel C. OCIO index and OCIO search consultant



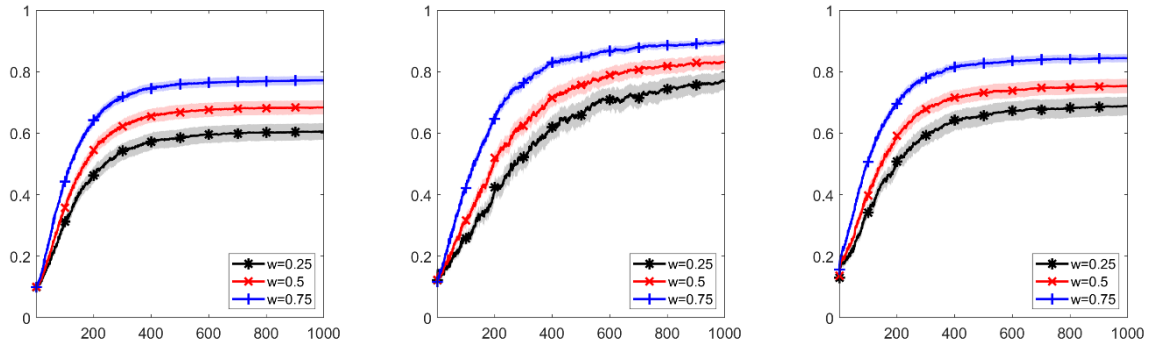
Note: This figure presents the differences in the optimal OCIO selection proportions between exploration methods based on changes in the number of managers when the weight on preference is fixed at 0.5 ($w = 0.5$). Panels A, B, and C present the differences between the OCIO index and self-directed searching, OCIO search consultants and self-directed searching, and OCIO index and OCIO search consultant, respectively. The vertical axis represents the proportion of asset owners who select the optimal OCIO, while the horizontal axis represents the iteration number. “Self,” “Index,” and “Consultant” refer to different exploration methods—specifically, self-directed searching, using the OCIO index, and consulting with an OCIO search consultant, respectively. Each line on the graph represents the average of 100 simulation results for a given scenario. The dotted line, dashed line, and solid line indicate the result for a scenario with 5, 10, and 20 managers, respectively. The shading around each line represents a 95% confidence interval.

Figure 4. Proportion of the optimal OCIO selection based on the changes in weight on preference
Panel A. Performance ($w = 0.25$) *Panel B.* Baseline scenario ($w = 0.5$) *Panel C.* Preference ($w = 0.75$)



Note: This figure presents the proportion of optimal OCIO selection based on changes in the weight on preference when the number of managers is fixed at 5 ($m = 5$). Panels A, B, and C present the results of scenarios wherein the weight is 0.25, 0.5, and 0.75, respectively. The vertical axis represents the proportion of asset owners who select the optimal OCIO, while the horizontal axis represents the iteration number. “Self,” “Index,” and “Consultant” refer to different methods of exploration—specifically, self-directed searching, using the OCIO index, and consulting with an OCIO search consultant, respectively. Each line on the graph represents the average of 100 simulation results for a given scenario. The lines with circle, triangle, and square markers indicate the results of scenarios using the self-directed searching method, the OCIO index method, and the OCIO search consultant method, respectively. The shading around each line represents a 95% confidence interval.

Figure 5. Comparison of the changes in the weight on preference by exploration method
Panel A. Self-directed searching *Panel B.* OCIO index *Panel C.* OCIO search consultant



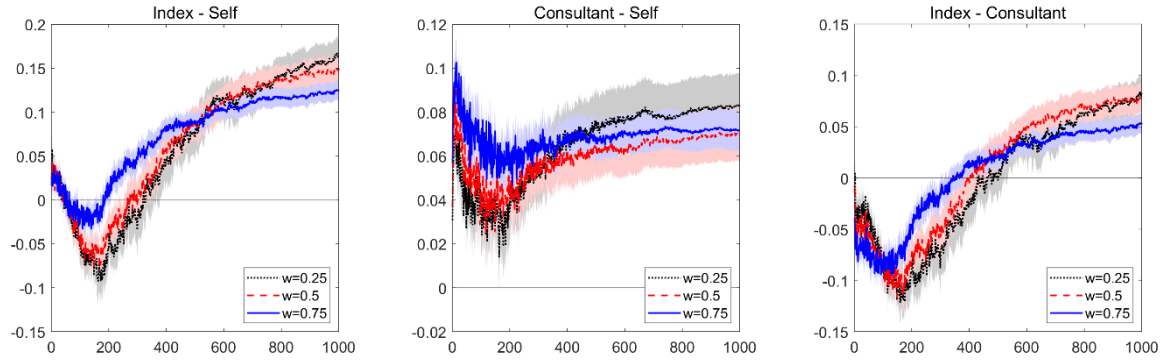
Note: This figure presents the comparison of changes in weight on preference by exploration method when the number of managers is fixed at 5 ($m = 5$). Panels A, B, and C present the results of scenarios with self-directed searching, the OCIO index, and the OCIO search consultant, respectively. The vertical axis represents the proportion of asset owners who select the optimal OCIO, while the horizontal axis represents the iteration number. Each line on the graph represents the average of 100 simulation results for a given scenario. The lines with asterisk, x-shaped, and cross markers indicate the results of scenarios with weights of 0.25, 0.5, and 0.75, respectively. The shading around each line represents a 95% confidence interval.

Figure 6. Differences in the optimal OCIO selection proportions between exploration methods based on changes in the weight of preference

Panel A. OCIO index and self-directed searching

Panel B. OCIO search consultant and self-directed searching

Panel C. OCIO index and OCIO search consultant



Note: This figure presents differences in optimal OCIO selection proportions between exploration methods based on changes in the weight of preference when the number of managers is fixed at 5 ($m = 5$). Panels A, B, and C present the differences between the OCIO index and self-directed searching, OCIO search consultants and self-directed searching, and OCIO index and the OCIO search consultant, respectively. The vertical axis represents the proportion of asset owners who select the optimal OCIO, while the horizontal axis represents the iteration number. “Self,” “Index,” and “Consultant” refer to different methods of exploration—specifically, self-directed searching, using the OCIO index, and consulting with an OCIO search consultant, respectively. Each line on the graph represents the average of 100 simulation results for a given scenario. The dotted, dashed, and solid lines indicate the results of scenarios with weights of 0.25, 0.5, and 0.75, respectively. The shading around each line represents a 95% confidence interval.

Table 1. Summary of Notation

Notation	Meaning	Notation	Meaning
i	Index of the asset owner. $i = 1, \dots, n.$	n	Number of asset owners
a	Index of the manager. $a = 1, \dots, m.$	m	Number of managers
t	Index of iteration. $t = 1, \dots, T.$	T	Number of iterations.
ψ_i	Preference parameter of asset owner i .	$u_{a,i}$	Asset owner i 's preference for manager a .
τ_a	Preference parameter of manager a .	$r_{a,t}$	Return of manager a 's portfolio at the t -th iteration.
μ_a	Expected return of manager a 's portfolio.	\tilde{z}_t	Random shock at the t -th iteration, which follows a standard normal distribution.
$R_{i,t}(a)$	Reward when asset owner i chooses manager a at time t .	w	Weight on preference
$E[R_{i,t}(a)]$	Expected reward when asset owner i chooses manager a at time t .	$Q_{i,t}(a)$	Q -value when asset owner i chooses manager a at the t -th iteration.