Idea Sharing and the Performance of Mutual Funds*

Julien Cujean[†]

Abstract: I develop an equilibrium model to explain why few mutual fund managers consistently outperform, even though many have strong informational advantages. The key ingredient is that managers obtain investment ideas through idea sharing. Idea sharing improves statistical significance of alpha through increased price informativeness. But it also causes better informed managers to take larger positions, which makes their alpha noisier—although a significant fraction of managers build strong informational advantages, statistical significance and persistence of alpha concentrate in underperforming funds. I argue that in-house development of ideas cannot explain these facts.

JEL classification: D82, D83, D85, G23.

Keywords: Idea Sharing; Performance; Rational-Expectations Equilibrium.

*I am particularly grateful to Ron Kaniel and an anonymous referee for their valuable suggestions. I also wish to thank Bernard Dumas, Pierre Collin-Dufresne, Darrell Duffie and Julien Hugonnier for their guidance on a previous version of this paper, which circulated under the title "The Social Dynamics of Performance". I have benefited from the comments of Daniel Andrei, Philippe Bachetta, Maria Cecilia Bustamante, Harjoat Bhamra, Francesco D'Acunto, Jérôme Detemple, Laurent Frésard, Amit Goyal, Clifton Green, Michael Hasler, Steve Heston, Pete Kyle, Chi Liao, Mark Loewenstein, Semyon Malamud, Rich Mathews, Antonio Mele, Erwan Morellec, Ilaria Piatti, Rémy Praz, Alberto Rossi, Shri Santosh, Nic Schaub, Norman Schuerhoff, Philip Valta, Russ Wermers, Xiang Zhang, Bart Zhou Yueshen, Haoxiang Zhu, and seminar participants at the WFA meeting, the Adam Smith Asset Pricing Workshop, the LAEF Conference in Santa Barbara, the EFA meeting in Oslo, the International Conference on Capital Markets at INSEAD, the Gerzensee Search and Matching Workshop, the 10th International Paris Finance Meeting, the DEEP macro workshop at UNIL, the EPFL Research seminar, the NFA meeting in Ottawa, the 17th SGF Conference, the Gerzensee Doctoral Workshop, the Latsis Symposium at ETHZ, Northwestern University, UNC-Chapel Hill, University of Maryland, University of Rochester, McGill University, University of Toronto, Johns Hopkins University, Boston University, HEC Paris, University of Iowa, UT Dallas, Georgia Tech, NHH, the University of Bern, the Geneva Finance Research Institute, the University of Zurich and the University of Lugano.

[†]CEPR and University of Bern, Institute for Financial Management, Office 214, Engehaldenstrasse 4, CH - 3012 Bern; +41 (0) 31 631 3853; julien.cujean@ifm.unibe.ch; webpage www.juliencujean.com.

1. Introduction

The vast majority of mutual fund managers are unable to generate abnormal returns, and a significant fraction underperforms passive benchmarks.[1] The handful of managers who outperform seldom maintain their performance. These stylized facts may suggest that managers do not possess superior information (e.g., Carhart (1997)). Yet, there exists evidence suggesting that managers do have informational advantages, which originate from superior ability to gather and analyze the data (Chevalier and Ellison, 1999), geographical proximity to investments (Coval and Moskowitz, 2001), education networks (Cohen, Frazzini, and Malloy, 2008) or knowledge spillovers in large cities (Christoffersen and Sarkissian, 2009).

I propose a mechanism that resolves this apparent contradiction. The central feature is that fund managers obtain investment ideas through idea sharing. A literal interpretation of idea sharing is that managers interact socially with each other (e.g., Hong, Kubik, and Stein (2005)).[2] A broader interpretation is that some managers purchase the same research from sell-side firms, heterogeneously sharing private information. Although idea sharing creates a rich heterogeneity of informational advantages across managers, it simultaneously leads the significance and persistence of performance to concentrate in underperforming funds.

I base my argument on a rational-expectations equilibrium model (e.g., He and Wang (1995)) in which a population of fund managers possess heterogeneous numbers of ideas. Managers collectively clear the market, but speculate individually on their informational advantage. Speculation on individual ideas is a zero-sum game that managers play against each other. In this zero-sum game, *skill* is defined as the distance between a manager's number of ideas and the cross-sectional average number of ideas. Because the "average

[1] Jensen (1968), Malkiel (1995), Gruber (1996) and Carhart (1997) find that most mutual funds fail to outperform, and often underperform, passive benchmarks, even before transactions costs. See Jones and Wermers (2011) for a literature review.

[2] Interpersonal communication plays an important role in institutional investors' decisions (Shiller and Pound, 1989). Fund managers exploit information through word-of-mouth communication (Hong et al., 2005) and share profitable ideas (Gray, Crawford, and Kern, 2012; Pool, Stoffman, and Yonker, 2015). For instance, David Einhorn (a "star" manager) acknowledges that "*sometimes an analyst generates the idea, sometimes other fund managers, a conference, or an idea dinner*". (This quote is taken from an interview that is available at `http://www.marketfolly.com/2012/03/david-einhorns-extensive-q-session-from.html`.)

manager" clears the market, a manager's *alpha* is the return she generates relative to the return of the average manager. In other words, managers who hold more ideas than the average manager create positive alpha at the expense of managers who hold fewer ideas.

In this paper managers and the empiricist who evaluates their performance face opposite inference problems. While managers attempt to infer fundamentals in real time knowing their own skill, the empiricist attempts to infer managers' skill *ex-post* controlling for fundamentals. Allowing the empiricist to control for all information that was value relevant over the sample period rules out biases in alpha measurement. The empiricist knows when it was historically optimal to go long or short and whether or not each manager traded appropriately. However, what she does not know is whether an appropriate trade reflects skill or luck. Trades contain an element of *luck*, the noise in managers' ideas. The empiricist does not observe skill and luck separately, but only observes a nonlinear combination of the two.

In this context alpha is an unbiased yet noisy measure of skill. To generate an additional unit of value, an additional idea requires a manager to take larger positions. As a result, skilled managers load more aggressively on skill but also on luck and thus generate a higher but noisier alpha. In the model this tension between a manager's skill and luck is captured by the ratio of the two, which I call the *skill-to-luck ratio*. Importantly, this ratio is concave in a manager's number of ideas—an additional idea produces alpha with decreasing returns to scale in terms of statistical significance.[3] Because skilled managers generate a noisier alpha, an additional idea improves their alpha $t-$statistic by less than the alpha it creates.

The way managers gather ideas entirely determines how alphas and their $t-$statistics vary across managers and over time. The main argument of the paper is that idea sharing distinctly explains the facts. I model idea sharing using the information percolation theory (Duffie, Malamud, and Manso, 2009) through which managers share ideas in bilateral meetings; an equivalent interpretation is that pairs of managers are selected randomly to observe

[3]Here, "scale" refers to the number of ideas, as opposed to the traditional notion of fund size (e.g., Chen, Hong, Huang, and Kubik (2004)). In the model, decreasing returns to scale in fund management are absent, because managers have CARA utility and trade competitively.

the same private information (e.g., they buy the same research). This framework extends Andrei and Cujean (2017) to continuous trading and a broader information structure.

Developing information in-house (e.g., Branikas, Hong, and Xu (2017)) is a natural alternative to idea sharing. Sharing ideas and originating ideas have distinct implications regarding how ideas flow over time and how they are distributed cross-sectionally. Suppose ideas arise at a fixed arrival rate, both when shared or originated. Under this assumption, the defining feature of idea sharing is that a manager's knowledge increases with knowledge possessed by other managers. Knowledge thus feeds upon itself, causing information to flow exponentially and a strong left skew in the cross-sectional distribution of ideas. When managers develop ideas on their own this feedback is absent; in the extreme, knowledge produced in-house is fixed. Information thus flows linearly and numbers of ideas are distributed symmetrically across managers. The relevance of idea sharing is that it can explain stylized facts about fund performance that idea origination cannot.

I first examine the conditions under which an empiricist identifies significant alpha. Idea sharing raises the statistical significance of alpha by increasing price informativeness. However, it simultaneously causes skilled managers to take larger positions, making their alpha noisier. Significant alphas thus exist under idea sharing, but concentrate in underperforming funds. Consistent with stylized facts, most managers are unable to generate abnormal returns, a significant fraction underperforms, whereas a handful of top performers spreads in the far right tail (e.g., Barras, Scaillet, and Wermers (2010)). In contrast, when ideas are originated, the empiricist fails to identify significant performance (negative or positive)—she even fails to reject the null hypothesis of zero alpha for a perfect market timer.

Similarly, the separation of skill from luck concentrates in underperforming funds when ideas are shared, whereas skill is virtually indistinguishable from luck when ideas are originated. In the model the performance of the average manager defines the null hypothesis of pure luck. The left skew in the distribution of ideas when shared causes the concavity of the skill-to-luck ratio to shift mass from the right to the left of the cross-sectional distribution of $t-$statistics, tilting it to the left relative to the distribution under the null. Hence, idea sharing can explain why most underperforming funds appear to be truly unskilled, whereas outperforming funds appear to be lucky (e.g., Fama and French (2010)). In contrast, because the distribution of ideas is symmetric when originated, the cross-sectional distribution and that under the null are indistinguishable.

To study performance persistence I segment the population of managers into groups that differ by quality of ideas. Whether ideas are shared or originated has similar implications for persistence (except asymptotically[4]); what matters is whether segmentation arises endogenously through networks, which allows alpha and its $t-$statistic to move in opposite directions. Because the skill-to-luck ratio is concave, a decline or an improvement in skill affects unskilled managers' performance comparatively more than it does for skilled managers. Network formation further leads skill to deteriorate at an increasing rate and to improve at a decaying rate, thus leading performance persistence to concentrate in underperforming funds (e.g., Carhart (1997)). Instead, when the population is segmented exogenously into groups of skilled and unskilled managers, alphas and their $t-$statistic converge across groups.

These results lead to new empirical implications. $T-$statistics have undesirable cross-sectional properties due to the concavity of the skill-to-luck ratio. Controlling for managers' trading activity may help reduce this effect (Pastor, Stambaugh, and Taylor, 2016). However, $t-$statistics may retain the ability to rank managers consistently even when associated alphas converge. The empiricist may also fail to reject the null hypothesis of zero alpha for a perfect market timer. To make the null more difficult to reject, the empiricist should determine the maximal level of statistical significance she expects to observe in sample.

The leading theory for the lack of performance persistence of mutual funds is based on fund flows—abnormal performance attracts fund flows that compete away subsequent abnormal returns (Berk and Green, 2004). I abstract from fund flows and provide an explanation for the absence of performance even when fund flows are irrelevant (e.g., closed-end funds

[4]Perfect market timing keeps delivering significant alpha asymptotically only when ideas are shared.

(Wu, Wermers, and Zechner, 2015)). Note, however, that incorporating fund flows leaves my results qualitatively unchanged. Fees are another important institutional feature. Because mutual funds fees are of the fulcrum type (Cuoco and Kaniel, 2011), they do not affect the results in my framework. Other theories of fund flows and fees include Spiegel and Mamaysky (2001), Basak, Pavlova, and Shapiro (2007), Glode (2011), Vayanos and Woolley (2011), Pastor and Stambaugh (2012) and Kaniel and Kondor (2013).

More broadly, like Admati and Ross (1985) and Dybvig and Ross (1985), I study performance in a rational-expectations framework.[5] However, I rule out measurement biases in alpha and focus on statistical significance. This paper also adds to the literature studying information percolation in centralized markets (e.g., Andrei and Cujean (2017)).[6] The novelty of this paper is to rationalize why alpha mostly detects underperformance.

In the remainder of the paper, all proofs are relegated to the Appendix. For convenience, Table 1 provides a summary of the notation used throughout the paper.

2. A model of fund managers gathering ideas

I start by describing a process through which managers collect investment ideas in chunks (Section 2.1). I then embed this process of idea gathering in an otherwise standard rational-expectations model—a continuous-time version of He and Wang (1995) (Section 2.2).

2.1. Building a collection of investment ideas in chunks

I consider a continuous-time economy with a finite horizon T, at which some unobservable dividend $\Pi \sim \mathscr{N}(0, \sigma_\Pi^2)$ will be paid. I refer to this liquidating dividend as the fundamental. The economy is populated with a continuum of fund managers indexed by $i \in [0, 1]$. Over

[5]See Grinblatt and Titman (1989), Kothari and Warner (2001), Goetzmann, Ingersoll, Spiegel, and Welch (2007) and Mamaysky, Spiegel, and Zhang (2008) for performance measures in alternative frameworks.

[6]Related works on social interactions use alternative information-sharing mechanisms. Some references use graph theory (Ozsoylev and Walden (2011), Colla and Mele (2010), Acemoglu, Bimpikis, and Ozdaglar (2010) or DeMarzo, Vayanos, and Zwiebel (2003)) and others use epidemiological models (Hong, Hong, and Ungureanu (2010) and Burnside, Eichenbaum, and Rebelo (2010)).

time each manager i obtains an increasing sequence of private signals about the fundamental:

$$S_j^i = \Pi + \epsilon_j^i, \quad j = 1, ..., n_t^i \tag{1}$$

where $n_t^i \in \mathbb{N}^*$ denotes the number of signals manager i has collected up to time t and where $\epsilon_j^i \sim \mathcal{N}(0, \sigma_S^2)$ represents the "idiosyncratic" noise in manager i's $j-$th signal. By *idiosyncratic* I mean that there is one such random variable per manager i and signal j, and that these random variables are sufficiently independent for a version of the Strong Law of Large Numbers to hold across managers and signals (e.g., Duffie and Sun (2007)).[7]

I view manager i's set of signals in Eq. (1) as a collection of investment ideas. I now describe the process through which managers build heterogeneous collections of ideas. A manager starts with an initial, idiosyncratic number n_0^i of ideas, which is drawn from a distribution π_0 with support $\mathbb{N}^*$. She then collects new ideas at arrival times of an idiosyncratic Poisson process $(N_t^i)_{t \geq 0}$ with time-varying intensity $\eta_t(n_{t-}^i)$. The intensity at which she gets new ideas potentially depends on her current number n_{t-}^i of ideas. For instance, a manager who has gathered many ideas may be more efficient at collecting new ideas in the future.

Whenever a manager gathers new ideas—say at time t—she receives a chunk $(S_{j+n_{t-}^i}^i : 1 \leq j \leq \Delta n_t^i)$ of ideas. The incremental number Δn_t^i of ideas is drawn from a distribution $\pi_t(\cdot; n_{t-}^i)$, which potentially depends on her current number n_{t-}^i of ideas. Since individual ideas are Gaussian and independent (conditional on Π), a manager's average new idea:

$$Y_t^i \equiv \tfrac{1}{\Delta n_t^i} \textstyle\sum_{j=1}^{\Delta n_t^i} S_{j+n_{t-}^i}^i = \Pi + \tfrac{\sigma_S}{\sqrt{\Delta n_t^i}} \epsilon_t^i, \quad \epsilon_t^i \sim \text{i.i.d. } \mathcal{N}(0,1) \tag{2}$$

is a sufficient statistic for the chunk of new ideas she receives. In other words, a manager's private information is completely summarized by two numbers at any time t, her average idea and her total number n_t^i of ideas, which represents the quality of her information.

This process of idea gathering generates a cross-sectional distribution of number of ideas,

[7] That is, for almost every pair (i, i') of managers ϵ_j^i and $\epsilon_{j'}^{i'}$ are pairwise independent for all j and j'.

which I denote by $\mu_t(n)$. This distribution keeps track of the number n of ideas across the population of managers at every date t; given the assumptions above, it satisfies the following differential equation (see Appendix F):

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\mu_t(n) = -\eta_t(n)\mu_t(n) + \textstyle\sum_{m=1}^{n-1} \eta_t(n-m)\mu_t(n-m)\pi_t(m; n-m), \quad \mu_0(n) = \pi_0(n). \tag{3}$$

The first term on the right-hand side in Eq. (3) is the rate at which managers leave their type, i.e., the fraction of managers of type n who received new ideas and thus no longer hold n ideas. The second term represents the rate at which managers enter a new type. If a manager of current type $n-k$ receives k new ideas, she becomes of type n.

A key statistic for the equilibrium analysis is the cross-sectional average number of ideas:

$$\phi_t \equiv \sum_{n\in\mathbb{N}} \mu_t(n)n, \tag{4}$$

the first moment of the cross-sectional distribution in Eq. (3). This average determines the rate at which private information flows in the population of managers. In Section 4, I discuss different interpretations of the process of idea gathering—different specifications of the intensity η and the distribution π—under which Eqs. (3) and (4) have simple solutions.

2.2. The economy

I insert the mechanism of idea collection of Section 2.1 in a continuous-time version of He and Wang (1995). The resulting framework extends Andrei and Cujean (2017) to continuous trading and a broader information structure, and is analytically simpler. Managers exhibit CARA utility over consumption with a common coefficient of absolute risk aversion, γ. Trading takes place continuously between time 0 and the horizon T. The market consists of two assets. The first asset is a risky stock with equilibrium price P_t at time t. The stock is a claim to the liquidating dividend Π, its fundamental value. The second asset is a riskless claim with perfectly elastic supply and a rate of return normalized to $r = 0$. Its rate is exogenous because consumption and stock payout take place only once (at the horizon T).

The problem of a manager i is to find a predictable portfolio strategy θ^i maximizing her expected utility over terminal wealth

$$\mathbb{E}\left[-e^{-\gamma W_T^i}\middle|\mathscr{F}_t^i\right] \tag{5}$$

$$\text{subject to} \quad W_T^i = W_0^i + \int_{[0,T)} \theta_t^i \mathrm{d}P_t + \theta_T^i \Delta P_T, \tag{6}$$

where a manager i's information set at time t,

$$\mathscr{F}_t^i = \sigma\left(\left(P_s, S_j^i\right) : 0 \leq s \leq t, 1 \leq j \leq n_t^i\right), \tag{7}$$

contains two sources of information, *(1)* her private collection of ideas, $(S_j^i : 1 \leq j \leq n_t^i)$, which she builds through the mechanism of the previous section, and *(2)* the history of prices, $(P_t)_{t\geq 0}$, which is endogenous and publicly available.

The budget constraint in Eq. (5) includes a price discontinuity of size ΔP_T occurring at the horizon date. Because managers are risk averse and trade competitively, they do not completely exhaust their informational advantage by the horizon date. As a result, the price right before the horizon date P_{T-} is not equal to its fundamental value Π on average, but only matches its fundamental value *at* the horizon date. In continuous time, this phenomenon results in a price discontinuity $\Delta P_T = \Pi - P_{T-}$ when the stock pays out. To emphasize the particularity of this result, notice that in a continuous-time Kyle model, the price converges to the fundamental on average (Back, 1992).

Furthermore, the budget constraint in Eq. (5) ignores flows in and out of the fund. Fund flows have well-known effects on performance (Berk and Green, 2004), which may interact with the effects of managers' flow of ideas. To isolate informational effects on managers' performance, I first abstract from fund flows. In Section 6 I subsequently extend the model to incorporate fund flows and fees and examine how both features along with information flows jointly affect managers' performance. Finally, for simplicity, I do not impose short-sales or borrowing constraints, although managers may face certain trading restrictions.[8]

To prevent prices from fully revealing the fundamental, I assume that the supply Θ of the stock is random and follows an Ornstein-Uhlenbeck process,

$$\mathrm{d}\Theta_t = -a_\Theta \Theta_t \mathrm{d}t + \sigma_\Theta \mathrm{d}B_t^\Theta, \quad \Theta_0 \sim \mathscr{N}\left(0, \sigma_\Theta^2\right) \tag{8}$$

with volatility $\sigma_\Theta > 0$ and where $(B_t^\Theta)_{t\geq 0}$ is a Brownian motion. As is customary, I interpret Θ as the supply of the stock available to the market, while noise traders—agents who trade for reasons unrelated to fundamental information—have inelastic demands of $1 - \Theta$ units of the stock (in total supply of 1). To avoid inducing artificial persistence in managers' performance through the supply, I restrict the supply to be a martingale (i.e., $a_\Theta \equiv 0$). Under this specification, the persistence of performance depends exclusively on the pattern of information arrival, thus isolating the link between the flow of ideas and performance.

3. Equilibrium with continuous-discrete learning

In this section I obtain the equilibrium solution, from which I infer the concepts of "skill" and "alpha" that matter in equilibrium. To construct the equilibrium, I conjecture a price function, which I use to solve managers' filtering problem (Section 3.1). I then solve their optimization problem (Section 3.2) and clear the market, which allows me to verify the price conjecture, to define skill in equilibrium and to derive the associated alpha (Section 3.3).

3.1. Learning tick by tick and in large, infrequent chunks

A novel aspect is the combination of continuous-time trading and the discontinuous process of ideas gathering, which produces a continuous-discrete form of learning. I denote by $\widehat{\Pi}_t^i \equiv \mathbb{E}\left[\Pi \middle| \mathscr{F}_t^i\right]$ a manager i's expectations of the fundamental, by $\widehat{\Theta}_t^i \equiv \mathbb{E}\left[\Theta_t \middle| \mathscr{F}_t^i\right]$ her

[8] See, for instance, Almazan, Brown, Carlson, and Chapman (2004).

expectations of the supply and by $o_t^i \equiv \mathbb{V}\left[\Pi|\,\mathscr{F}_t^i\right]$ her posterior variance, each conditional on her information set $\mathscr{F}_t^i$ at time t. Similarly, I denote common expectations by $\widehat{\Pi}_t^c \equiv \mathbb{E}\left[\Pi|\,\mathscr{F}_t^c\right]$ and common uncertainty by $o_t^c \equiv \mathbb{V}\left[\Pi|\,\mathscr{F}_t^c\right]$, each conditional on the common information set, $\mathscr{F}_t^c = \sigma(P_s : 0 \leq s \leq t)$, which contains the (commonly observable) history of prices.

I focus the analysis on a linear equilibrium in which the stock price is a linear function of the state variables of the economy. Specifically, I conjecture that the price satisfies

$$P_t = \lambda_{1,t}\Pi + (1-\lambda_{1,t})\widehat{\Pi}_t^c + \lambda_{2,t}\Theta_t, \quad \forall t < T, \tag{9}$$

where $\lambda_{1,t}$ and $\lambda_{2,t}$ are deterministic functions to be solved for in equilibrium. Based on this conjecture, a manager i's expectations and common expectations evolve according to the dynamics that I highlight in Proposition 1.

Proposition 1. *In a linear equilibrium, conditional expectations solve the filtering equations*

$$d\begin{pmatrix} \widehat{\Pi}_t^c \\ \widehat{\Theta}_t^c \end{pmatrix} = \begin{pmatrix} o_t^c k_t \\ \sigma_\Theta - \frac{\lambda_{1,t}}{\lambda_{2,t}} o_t^c k_t \end{pmatrix} d\widehat{B}_t^c, \tag{10}$$

$$d\begin{pmatrix} \widehat{\Pi}_t^i \\ \widehat{\Theta}_t^i \end{pmatrix} = \begin{pmatrix} o_t(n_{t-}^i) k_t \\ \sigma_\Theta - \frac{\lambda_{1,t}}{\lambda_{2,t}} o_t(n_{t-}^i) k_t \end{pmatrix} d\widehat{B}_t^i + \begin{pmatrix} 1 \\ -\frac{\lambda_{1,t}}{\lambda_{2,t}} \end{pmatrix} \frac{o_t(n_t^i)\Delta n_t^i}{\sigma_S^2} \widehat{Y}_t^i\, dN_t^i, \tag{11}$$

with initial conditions given by

$$\begin{pmatrix} \widehat{\Pi}_0^c \\ \widehat{\Theta}_0^c \end{pmatrix} = \frac{\lambda_{1,0}\Pi + \lambda_{2,0}\Theta_0}{\lambda_{1,0}^2\sigma_\Pi^2 + \lambda_{2,0}^2\sigma_\Theta^2} \begin{pmatrix} \lambda_{1,0}\sigma_\Pi^2 \\ \lambda_{2,0}\sigma_\Theta^2 \end{pmatrix} \quad \textit{and} \quad \begin{pmatrix} \widehat{\Pi}_0^i \\ \widehat{\Theta}_0^i \end{pmatrix} = \begin{pmatrix} o_0(n_0^i)\left(\frac{\widehat{\Pi}_0^c}{o_0^c} + \frac{n_0^i Y_0^i}{\sigma_S^2}\right) \\ \widehat{\Theta}_0^c - \frac{\lambda_{1,0}}{\lambda_{2,0}}\left(\widehat{\Pi}_0^i - \widehat{\Pi}_0^c\right) \end{pmatrix} \tag{12}$$

where k, the speed at which prices reveal information, and conditional variances satisfy

$$k_t = \frac{1}{\sigma_\Theta}\frac{d}{dt}\frac{\lambda_{1,t}}{\lambda_{2,t}}, \quad o_t^c = \left(\frac{1}{\sigma_\Pi^2} + \left(\frac{\lambda_{1,0}}{\lambda_{2,0}}\right)^2 \frac{1}{\sigma_\Theta^2} + \int_0^t k_s^2\, ds\right)^{-1} \quad \textit{and} \quad o_t(n) = \left(\frac{1}{o_t^c} + \frac{n}{\sigma_S^2}\right)^{-1}$$

and where the process $(\widehat{B}^i_t)_{t\geq 0}$ *is a Brownian motion with respect to* $\mathscr{F}^i_t$, $\widehat{Y}^i_t \sim \mathscr{N}\left(0, o_t(n^i_{t-}) + \frac{\sigma_S^2}{\Delta n^i_t}\right)$ *is Gaussian conditional on* $\mathscr{F}^i_{t-}$ *and* Δn^i_t*, the process* $(\widehat{B}^c_t)_{t\geq 0}$ *is a Brownian motion with respect to* $\mathscr{F}^c_t$*, and* $(N^i_t)_{t\geq 0}$ *is a Poisson process with rate* $\eta_t(n)$*.*

A manager's i expectations in (11) are continuous-discrete, reflecting the pattern of information arrival associated with gathering ideas in large, infrequent chunks. In contrast, common expectations in (10) are continuous. In Figure 1 I illustrate this specific feature of the model with a simulation of a manager's number of ideas (left panel), her posterior variance (the center panel) and her expectations of the fundamental (the right panel).

[insert Figure 1 here]

As long as a manager does not get new ideas (the flat parts in the left panel), she extracts information from prices in an effort to infer ideas collected by other managers in the market. This information flows tick by tick, producing continuous updates in her expectations (the right panel) and her variance (the center panel). When a manager collects new ideas (the steep parts in the left panel), she gets information that comes in large pieces at discrete random times, leading to vast revisions in her expectations and substantial improvement in her precision. Continuous trading allows prices to reflect new ideas instantly—was trading discrete, prices would ignore them until the next trading round (Andrei and Cujean, 2017).

3.2. Optimal portfolio strategy

A difficulty in deriving a manager's portfolio is that she cannot hedge the risk associated with new ideas, because she cannot trade claims that pay off exactly when these ideas arise. She thus revises her expectations in a way that her marginal utility jumps after getting new ideas, which in principle precludes a linear equilibrium.[9] This issue is absent when trading is discrete (Andrei and Cujean, 2017), since new ideas cannot be exploited until the next trading

[9] In general, a framework involving jump-diffusion state variables, as those of Proposition 1, is of the linear-quadratic class only if jumps are restricted to linear processes (Cheng and Scaillet, 2007).

date. Remarkably, however, because managers are Bayesian they rationally anticipate these revisions and their portfolio therefore remains linear, the result of Proposition 2.

Proposition 2. *A manager i builds an optimal portfolio strategy θ_t^i at time t given by*

$$\theta_t^i \equiv \theta_t(\widehat{\Pi}^i - \widehat{\Pi}^c, \widehat{\Theta}^i, n^i) = \frac{A_{Q,t} - B_{Q,t}\left(o_t(n^i)\right)^{-1} B_{\Psi,t}(n^i)^\top \Lambda_t}{\gamma B_{Q,t}^2} \begin{pmatrix} \widehat{\Pi}^i - \widehat{\Pi}^c \\ \widehat{\Theta}^i \end{pmatrix} \tag{13}$$

where A_Q, B_Ψ and Λ are deterministic matrices and B_Q is a deterministic scalar.

A well-known result is that an increase in information precision stimulates trading (Grossman and Stiglitz, 1980). This effect arises through a manager's hedging demand (the second term in (13)), which scales with her precision, $\left(o_t(n^i)\right)^{-1}$: the more ideas she has, the larger the position she takes. A new aspect associated with gathering ideas in chunks, however, is the rate at which it stimulates trading. The literature commonly assumes the precision of private information increases linearly over time (e.g., Brennan and Cao (1997)). Since hedging demands scale with precision, average market trading inherits this linear increase. In this paper, as in Andrei and Cujean (2017), average market trading increases with the cross-sectional average number of ideas ϕ in Eq. (4), which may not be linear in time.

As is customary in the literature (e.g., He and Wang (1995)), a manager's portfolio reflects a balance between two trading motives. Writing the portfolio in (13) as (see Appendix D):

$$\theta_t^i \equiv \underbrace{d_{\Delta,t}(n^i)\overbrace{\left(1 - \left(\frac{o_t^c}{o_t(n^i)}\right)^{-1}\right)\left(\frac{1}{n_t^i}\sum_{j=1}^{n_t^i} S_j^i - \widehat{\Pi}_t^c\right)}^{\widehat{\Pi}_t^i - \widehat{\Pi}_t^c \text{ (informational advantage)}}}_{\text{Speculative Position}} + \underbrace{d_{\Theta,t}(n^i)\overbrace{\left(\Theta_t - \frac{\lambda_{1,t}}{\lambda_{2,t}}(\widehat{\Pi}_t^i - \Pi)\right)}^{\widehat{\Theta}_t^i \text{ (perceived supply)}}}_{\text{Market-Making Position}} \tag{14}$$

shows that a manager speculates on her informational advantage and accommodates noise trading. The extent to which a manager speculates on her informational advantage depends on her precision relative to common precision, o^c/o. Her market-making activity depends on her perception of the supply. Unless a manager has infinitely many ideas, she is imperfectly informed and may thus mistakenly interpret an informed trade (as measured by her informational disadvantage $\widehat{\Pi}^i - \Pi$, the second term in (14)) as a noninformational trade, Θ. Depending on her number of ideas, she adjusts her market-making position to provide liquidity or to minimize the chance that she ends up on the wrong side of the trade.

Using the portfolio decomposition in (14), I introduce the concept of informational holding (He and Wang, 1995; Brennan and Cao, 1997), which I use throughout the analysis.

Definition 1. *A manager's informational holding $\widehat{\theta^i} := \theta^i - \Theta$ is a manager's total position net of per capita supply shock.*

When information is homogeneously perfect, every manager takes a position $\theta^i = \Theta$ that exclusively reflects noninformational trading. Therefore, a manager's informational holding $\widehat{\theta^i}$ identifies the part of her position that is purely generated by differential information, relative to her position under perfect information. Unless a manager has infinitely many ideas, she does not observe the supply and thus her own informational holding. However, the econometrician, who observes the portfolio of all managers *ex-post*, can compute informational holdings retrospectively. I elaborate on this aspect in Section 3.3.3.

3.3. Equilibrium

Combining Proposition 2 and Definition 1, the market-clearing condition at time t satisfies

$$\underbrace{\int_0^1 \widehat{\theta_t^i} \mathrm{d}i}_{\text{Informational Trading}} = 0. \tag{15}$$

I use the market-clearing condition in (15) to compute the price coefficients in (9) and to define a zero-sum game among managers. I then infer the definition of skill that is relevant in equilibrium and compute the alpha that managers generate on differential information.

3.3.1. Equilibrium solution and price informativeness

In equilibrium, as in Andrei and Cujean (2017), the cross-sectional average number of ideas ϕ in Eq. (4) completely determines price coefficients and the speed of information revelation, as Proposition 3 demonstrates.

Proposition 3. *In the unique linear equilibrium of the economy, the price coefficients satisfy*

$$\lambda_{1,t} = \frac{\phi_t o_t^c}{\sigma_S^2 + \phi_t o_t^c} \quad \textit{and} \quad \lambda_{2,t} = -\gamma \frac{o_t^c \sigma_S^2}{\sigma_S^2 + \phi_t o_t^c}, \quad t < T. \tag{16}$$

Hence, prices reveal information at speed $k_t = -\frac{1}{\sigma_\Theta \sigma_S^2 \gamma} \frac{d}{dt}\phi_t$ *and common uncertainty satisfies*

$$o_t^c = \left(\frac{1}{\sigma_\Pi^2} + \left(\frac{\phi_0}{\gamma \sigma_\Theta \sigma_S^2} \right)^2 + \left(\frac{1}{\sigma_\Theta \sigma_S^2 \gamma} \right)^2 \int_0^t \left(\frac{d}{ds}\phi_s \right)^2 ds \right)^{-1}. \tag{17}$$

Price informativeness improves with the precision of private information (Grossman and Stiglitz, 1980)—the average flow of ideas ϕ. Price informativeness is measured by the signal-noise ratio, which using (16) satisfies $\frac{\lambda_1}{\lambda_2} = -\frac{\phi}{\gamma \sigma_S^2}$. The speed at which prices reveal information is thus proportional to the rate at which the cross-sectional average number of ideas increases over time. An increase in this average stimulates trading, which leads to an increase in the signal-noise ratio and in common precision (the inverse of common uncertainty in (17)). Hence, price informativeness is measured in units of average number of ideas.

3.3.2. Skill in equilibrium: informational trading as a zero-sum game

Market making is a game that managers collectively play against noise traders. In contrast, trading based on differential information—informational trading—is a zero-sum game that managers play against each other, as the market-clearing condition in (15) indicates. Because managers' information is noisy, informational trading (see Definition 1) can be decomposed into an information component and a noise component, as in Proposition 4.

Proposition 4. *The informational holding $\widehat{\theta^i}$ of manager i admits the decomposition*

$$\widehat{\theta^i_t} = \underbrace{\frac{n^i_t - \phi_t}{\gamma\sigma^2_S|k_t|} \times \overbrace{\frac{\sigma^2_S|k_t|\left(\Delta_t + \gamma o^c_t \Theta_t\right)}{\sigma^2_S + o^c_t\phi_t}}^{\substack{\textit{Sharpe Ratio } \equiv SR_t \\ \textit{(under perfect information)}}}}_{\textit{information (skill)}} + \underbrace{\frac{\sqrt{n^i_t}}{\gamma\sigma_S}}_{\textit{noise (luck)}}\epsilon^i_t \tag{18}$$

where $\Delta_t \equiv \Pi - \widehat{\Pi}^c_t$ is the informational advantage achieved under perfect information.

The decomposition in (18) isolates the extent to which a manager's trades reflect skill or luck. Consider first a manager who has infinitely many ideas. Since this manager has a perfect informational advantage, Δ, and a perfect knowledge of the supply, Θ, her trades are purely driven by skill (information): depending on which component, Δ or Θ, drives the Sharpe ratio, SR, in (18), she trades towards the fundamental, Π, and against the common opinion, $\widehat{\Pi}^c$, or towards noise traders' demand. Consider now the "average manager" who holds the cross-sectional average number $n^i = \phi$ of ideas. Her informational holdings satisfy:

$$\widehat{\theta^i_t} = \frac{\sqrt{\phi_t}}{\gamma\sigma_S}\epsilon^i_t. \tag{19}$$

That is, her trades are purely driven by luck (noise). If the Sharpe ratio, SR, is high—either because the common opinion underestimates the fundamental ($\Delta > 0$) or noise traders sell ($\Theta > 0$), or both—and she happens to buy ($\widehat{\theta^i} > 0$), then she had a lucky draw ($\epsilon^i > 0$).

In general, the decomposition in (18) implies that skill is defined as follows.

Definition 2. *A manager's skill $s_t(n^i) := n^i - \phi_t$ is the distance between her number of ideas and the cross-sectional average (i.e., her precision relative to that of the average manager).*

Since the average manager is purely lucky, all managers who hold fewer ideas than ϕ trade against the perfectly informed manager on average, while all others are truly skilled.

3.3.3. Informational alphas

To construct a measure of alpha, I adopt the perspective of the econometrician. The main idea is that the empiricist and managers face opposite inference problems. Whereas managers attempt to infer fundamentals in real time knowing their own skill, the empiricist attempts to infer managers' skill *ex-post* controlling for fundamentals.

In this context alpha is an unbiased but noisy measure of skill. Assuming that the econometrician's dataset contains all information that was useful for portfolio management at each date of the sample rules out biases in alpha measurement. Unlike an uninformed observer (Dybvig and Ross, 1985), the empiricist controls for all relevant factors affecting the fundamentals of the economy. She knows when it was historically optimal to go long or short the stock and whether or not each manager traded appropriately. However, what the empiricist does not know is whether an appropriate trade reflects skill or luck, because she does not observe a manager's skill and luck separately; she only observes a manager's informational holding, which reveals a nonlinear combination of the two (see Proposition 4).

Formally, the empiricist observes past prices as well as the time series of returns $(\theta_t^i \mathrm{d}P_t)_{t \geq 0}$ that each manager i generates over the sample period $[0, T)$. This information is observed *ex-post*, when it is no longer value relevant to managers. Since price changes are continuous, the empiricist also observes the portfolio of each manager. She then obtains the supply at each date by averaging portfolios across managers, and recovers the informational holding $\widehat{\theta}^i$ of each manager i and the fundamental value by inverting the price. Thus, her dataset is:

$$\mathscr{F}_t = \sigma\left(\left(\Theta_s, \widehat{\theta}_s^i\right) : 0 \leq s \leq t, i \in [0, 1]\right) \bigvee \sigma(\Pi), \quad 0 \leq t < T. \tag{20}$$

Empiricists commonly use performance regressions to measure managers' performance. Following Koijen (2012), I express a manager's informational returns—returns she generates on her informational holding—as a performance regression formulated in continuous-time:

$$\widehat{\theta}_t^i \mathrm{d}P_t = \alpha_t^i \mathrm{d}t + \sigma_t^i \mathrm{d}B_t^{\Theta}, \tag{21}$$

in which informational returns, $\widehat{\theta}_t^i \mathrm{d}P_t$, are regressed on a constant, $\alpha_t^i \mathrm{d}t$. The residual noise, $\sigma_t^i \mathrm{d}B_t^{\Theta}$, is a Brownian motion. Evaluating managers based on informational returns allows me to focus the performance analysis on the zero-sum game of Section 3.3.2.

The empiricist uses all data $\mathscr{F}_t$ available up until time t to infer a manager's current alpha at time t (i.e., there is no look-ahead bias in estimating Eq. (21)). She obtains a time series of conditional alpha, α_t^i, and its $t-$statistic, $t_{\alpha,t}^i$, for each manager:

$$\alpha_t^i = \tfrac{1}{\mathrm{d}t}\widehat{\theta}_t^i \mathbb{E}[\mathrm{d}P_t|\mathscr{F}_t] \quad \text{and} \quad t_{\alpha,t}^i = \tfrac{\alpha_t^i}{\sigma_t^i} = \mathrm{sign}(\widehat{\theta}_t^i)\mathbb{E}[\mathrm{SR}_t|\mathscr{F}_t]. \tag{22}$$

I follow the convention of measuring alpha in dollar return (e.g., Dybvig and Ross (1985)). Note, however, that measuring alpha in rate of return leaves its $t-$statistic unaffected.[10]

On average a manager with n ideas is assigned the alpha and $t-$stat of Proposition 5.

Proposition 5. *The unconditional alpha estimate, $\widehat{\alpha}_t^i$, for a manager holding n^i ideas is:*

$$\widehat{\alpha}_t(n^i) = s_t(n^i)\frac{(|k_t| + \gamma\sigma_\Theta)o_t^c}{\gamma|k_t|(\sigma_S^2 + o_t^c\phi_t)}\mathbb{E}[SR_t^2] \tag{23}$$

at time t. Its unconditional $t-$statistic, $\widehat{t}_{\alpha,t}^i$, is (see Definition 3 below for $R_t(n)$):

$$\widehat{t}_{\alpha,t}(n^i) = sign(s_t(n^i))\sqrt{\frac{2}{\pi}}\mathbb{E}[SR_t^2]\Big(\mathbb{E}[SR_t^2] + R_t(n^i)^{-2}\Big)^{-\frac{1}{2}} \tag{24}$$

at time t. The unconditional squared Sharpe ratio in Eqs. (23) *and* (24) *satisfies:*

$$\mathbb{E}[SR_t^2] = k_t^2\left(o_t + o_t^2(\phi_t/\sigma_S^2 + \gamma^2\sigma_\Theta^2(t+1))\right) \tag{25}$$

[10]Replacing dollar return on the left-hand side of Eq. (21) with rate of return, $\frac{\widehat{\theta}_t^i}{|W_t^i|}\mathrm{d}P_t$, (i.e., dollar returns divided by assets under management) also leads to the $t-$statistic in Eq. (22).

where o^{-1} denotes average market precision, i.e., the precision of the average manager:

$$o_t^{-1} \equiv \int_0^1 (o_t^i)^{-1} di = \frac{1}{o_t^c} + \frac{\phi_t}{\sigma_S^2}. \tag{26}$$

By construction, unconditional alpha in (23) is an unbiased estimate of a manager's skill, s (see Definition 2). It evaluates a manager's ability to extract market-timing gains, as measured by the square of the Sharpe ratio (Treynor and Mazuy, 1966; Admati, Bhattacharya, Pfleiderer, and Ross, 1986). A manager appears to the empiricist as a successful market timer only if she is better informed than the average manager. Since skill is linear in number of ideas, an additional idea translates linearly into an additional unit of alpha in (23).

Although unbiased, a manager's informational alpha is a *noisy* estimate of her skill. To generate an additional unit of value, an additional idea requires a manager to take larger positions. Doing so, a manager not only increases her exposure to fundamental risk, but also to the noise in her average idea. An additional idea in turn has decreasing returns to scale in terms of Sharpe ratio and thus in terms of alpha $t-$statistic. In particular, the $t-$statistic in (24) depends on a manager's skill-to-luck ratio, R, which is defined as follows.

Definition 3. *A manager's skill-to-luck ratio, $R_t(n^i) := \frac{1}{\sigma_S|k_t|}\frac{s_t(n^i)}{\sqrt{n^i}}$, is the ratio of a manager's informational trading intensity on fundamental information relative to noise.*

This ratio is concave in a manager's number of ideas: because skilled managers generate a noisier alpha, an additional idea improves their $t-$statistic by less than the value it creates. In other words, whereas luck does not affect unconditional alphas, it affects their statistical significance. For this reason, alpha $t-$statistics are the focus of the analysis.

4. Economic interpretations of gathering ideas

The flow of private information, as measured by the cross-sectional average number of ideas, is central to the equilibrium analysis; it drives the concavity of the skill-to-luck ratio (see Definition 3), which itself drives the statistical significance of alpha. Thus, an intermediate step towards explaining stylized facts about alpha is to examine how different mechanisms of idea collection generate different patterns of information flow.

I consider two main sources from which managers collect ideas. First, mutual fund managers share ideas (e.g., Pool et al. (2015)). A literal interpretation of idea sharing is that managers interact socially with each other (e.g., Hong et al. (2005)). A broader interpretation is that managers purchase the same research from sell-side firms, thus sharing private information. Second, mutual funds spend considerable amounts of money to develop information in-house (e.g., Branikas et al. (2017)). These two mechanisms of idea collection—sharing ideas and originating ideas—imply different cross-sectional distributions of number of ideas and different flows of private information.

I model idea sharing using the information percolation theory (Duffie et al., 2009) through which managers share ideas in bilateral meetings; an equivalent but broader interpretation is that pairs of managers are selected randomly to observe the same private information (e.g., they buy the same research). Because each manager is a grain of sand in an ocean, for every such pair of managers, the sets of other managers with whom they respectively share ideas never overlap, nor does the same pair of managers ever meet twice. Thus, there is no sense of commonality of ideas—a manager learns an equivalent amount of information whether she gets an existing idea shared with another manager or whether she originates an idea with the same informational content. In other words, in this model, sharing existing information and producing original information are informationally equivalent.

Assuming ideas arise with fixed intensity η, both when shared or originated, I use the distribution π of incremental number of ideas to distinguish the two mechanisms. The essence of idea sharing is that knowledge obtained through idea sharing increases with knowledge possessed by other managers, which creates an endogenous relation between the distribution π and the cross-sectional distribution μ of number of ideas. The process of idea gathering thus feeds upon itself, leading information to flow exponentially. In contrast, when managers develop ideas on their own, the cross-sectional distribution of knowledge does not feed back in the knowledge they produce—the distribution π is fixed. Thus, information flows linearly.

4.1. Idea sharing among fund managers

I allow managers to share ideas according to the information percolation theory. Percolation is a matching mechanism whereby pairs of managers receive identical ideas; it can be interpreted indifferently as a meeting process through which managers exchange ideas or as a channel through which pairs of managers observe identical research. A key property of this matching process (to be described) is that it is additive in managers' number of ideas: two managers who have $n - m$ and m ideas, respectively, both wind up with n ideas after they are matched. As a result, the setup of Section 2.1 applies under this matching process—a manager's private information is summarized by her number of ideas and her average idea.

For simplicity I assume that all managers are initially endowed with exactly one idea at date $t = 0$. Thus, the initial distribution $\pi_0(n) = \delta_{n=1}$ of number of ideas has 100% mass at $n = 1$, where δ denotes the Dirac delta function. Whereas all managers start out with identical precisions, they can improve the precision of their ideas heterogeneously over time. From date $t = 0$ onward, a manager i is matched to other managers at arrival times of the Poisson process $(N_t^i)_{t\geq 0}$ of Section 2.1 with fixed intensity, η.

When two managers i and j are matched, their particular pairing is randomly sampled from the continuum of managers (Duffie and Sun, 2007). As is customary in the information percolation literature, the respective sets of other managers with whom managers i and j share ideas never overlap. In particular, because i and j belong to an infinite crowd of managers, they will not be matched again. While convenient from a modeling viewpoint, this property may sound unappealing in the context of fund management. To address this modeling caveat, in Section 5.2.2 I introduce bonds between managers through networks.

Once managers i and j are matched, they observe their respective collection of ideas in (1). Under the interpretation that managers actually meet, this assumption implies that they truthfully exchange their ideas, consistent with the fact that managers share valuable ideas (e.g., Pool et al. (2015)). Since managers i and j are infinitesimally small, they have no incentives to conceal, misrepresent nor add noise to their information—if they attempt to lie, they will not be able to individually influence prices and thus cannot benefit from their lies.[11] As anticipated above, the resulting matching process is additive in number of ideas.

This matching process requires that the distribution π of incremental number of ideas coincides with the cross-sectional distribution μ. Substituting the resulting identity $\pi \equiv \mu$ with $\eta_t(n) \equiv \eta$ into Eq. (3) yields the population dynamics in Andrei and Cujean (2017).

Corollary 1. *Under the idea-sharing mechanism, the cross-sectional distribution μ of ideas and the cross-sectional average number ϕ of ideas satisfy*

$$\mu_t(n) = e^{-\eta n t} \left(e^{\eta t} - 1\right)^{n-1} \quad \textit{and} \quad \phi_t = e^{\eta t} \ . \tag{27}$$

As Andrei and Cujean (2017) point out, in this setup the information flow is exponential.

4.2. Origination of ideas

Similar to idea sharing, I assume that managers develop ideas at fixed rate η. However, with idea origination the knowledge of others does not feed back in the knowledge managers produce independently. In the extreme, knowledge produced in-house is time invariant: $\pi_t = \pi_0$ at all times t. Under idea origination, π_0 is exogenous and thus arbitrary. But since I have assumed that all managers are initially endowed with one idea, it is natural to mirror this assumption, $\pi_0 \equiv \delta_{n=1}$. In other words, managers generate one idea at a time. In this context, the population dynamics in Eq. (3) take a particularly simple form.

Corollary 2. *Under the idea-origination mechanism, the cross-sectional distribution μ and*

[11] While infinitesimal managers do not have strict incentives to tell the truth, a finite number of managers may. Possible incentives include short-term investment horizons (Schmidt, 2015), reputation (Benabou and Laroque, 1992), or complementarity in information sets (Stein, 2008), for instance.

the cross-sectional average number ϕ of ideas satisfy

$$\mu_t(n) = e^{-\eta t}\frac{(\eta t)^{n-1}}{(n-1)!} \quad \textit{and} \quad \phi_t = \eta t + 1 \ . \tag{28}$$

The information flow under idea origination maps into the Brownian flow of signals that the literature commonly adopts (e.g., Detemple and Kihlstrom (1987)):

$$\mathrm{d}S_t^i = \Pi \mathrm{d}t + \sigma_S \mathrm{d}B_t^i, \quad S_0^i = \Pi + \sigma_S \epsilon_0^i \tag{29}$$

where $(B_t^i)_{t\geq 0}$ is an idiosyncratic Brownian motion. At the manager level, a collection of ideas developed in-house and the Brownian flow in Eq. (29) differ in several dimensions.[12] At the population level, however, both imply that the clock of information arrival (ϕ_t in Eq. (28), with $\eta \equiv 1$ when applied to Eq. (29)) is measured in affine units of calendar time.

Idea origination is a natural benchmark against which to compare idea sharing. Under idea sharing the clock of informational arrival is measured in exponential units of calendar time (see Corollary 1). Hence, while the clock of information arrival and calendar time coincide when ideas are originated, they do not when ideas are shared. This difference has specific implications for performance, which I now turn to.

5. Mapping performance into the flow of ideas

I revisit three stylized facts in the context of my model—few managers outperform (e.g., Barras et al. (2010)), their performance is indistinguishable from luck (e.g., Fama and French (2010)) and does not persist (e.g., Carhart (1997)). A unique implication of idea sharing is that statistical significance, the separation of skill from luck (Section 5.1) and performance persistence (Section 5.2) concentrate in the worst-performing funds. These facts are challenging to explain under idea origination.

[12] Ideas accumulate by chunks and come in heterogeneous numbers when originated (they are Poisson distributed), but flow in small, frequent pieces and have homogenous precision under Eq. (29).

5.1. Cross-sectional implications

I first analyze how idea sharing and idea origination generate distinct cross-sectional implications for the statistical significance of alphas and the separation of skill from luck.

5.1.1. Statistical significance of unconditional alpha

On average (unconditionally), at each date t the econometrician assigns to a manager i an alpha $t-$statistic, $\widehat{t}^i_{\alpha,t}$, according to Proposition 5. The resulting cross-sectional distribution of $t-$statistics thus evolves over time. To eliminate the effect of time when illustrating cross-sectional implications, I compute the average alpha $t-$statistic,

$$\overline{t}^i_{\alpha,T}(k) = \frac{1}{T}\int_0^T \widehat{t}_{\alpha,t}\left(\mathbb{E}\left[n^i_t \middle| n^i_T = k\right]\right)\mathrm{d}t, \tag{30}$$

over the trading period $[0, T)$.[13] The only source of cross-sectional variation across average alpha $t-$statistics in Eq. (30) is the (discrete) number k of ideas that a manager holds at the horizon date. Hence, the distribution of unconditional $t-$statistics in Eq. (30) is discrete, and follows from Corollary 1 or 2 depending on the mechanism of idea gathering considered.

I start with the following observation, which I formulate as a corollary.

Corollary 3. *In this framework without idea sharing or idea origination (i.e., $\eta \equiv 0$), all managers generate no informational alpha.*

Proof. Setting $\eta \equiv 0$, $n^i \equiv 1$ for all $i \in [0, 1]$ and thus, using (23), $\widehat{\alpha}^i_t \equiv 0$ for all $i \in [0, 1]$. □

In this idealized framework managers can only generate new ideas through sharing or in-house development; if managers are inactive or uncreative, they have homogeneous skill and thus no informational advantage over one another. That is, they only trade against noise traders, but not against each other. Market clearing then implies that they do not generate an informational alpha. What sharing ideas or originating ideas precisely do is to

[13] I assume that a manager i holds the average trajectory, $\mathbb{E}\left[n^i_t \middle| n^i_T = k\right]$, of number of ideas that leads her to hold k ideas by the horizon date. I compute this average trajectory in Appendix F.

create heterogeneity in skill. This heterogeneity causes managers to trade against each other, producing a cross-sectional distribution of alphas around the average manager.

[insert Figure 2 here]

Assessing statistical significance of managers' alphas requires computing the distribution of their $t-$statistics, which I plot in Figure 2. Each figure in this paper relies on an illustrative calibration, which I present in its description. In both panels of Figure 2 I indicate in red the average manager who generates no alpha by virtue of market clearing. Each bar represents the probability mass of each $t-$statistic. The main determinant of the distance between these bars is the concavity of the skill-to-luck ratio (see Definition 3).

The distributions of alpha $t-$statistics under idea sharing and idea origination differ along two dimensions. First, idea origination creates a distribution of $t-$statistics that is symmetric around the average manager, with most $t-$statistics spreading at equidistant intervals. In contrast, idea sharing creates a strong left skew in the distribution, making extreme outcomes on the left more likely; it also implies an asymmetry in the dispersion of $t-$statistics. This asymmetry results from the concavity of managers' skill-to-luck ratio, which is strongest under idea sharing due to the exponential flow of ideas it implies (see Corollary 1). Because managers in the left tail are unskilled, they take smaller positions, their alpha is less noisy and thus $t-$statistics spread over large intervals; because managers in the right tail are skilled, they take larger positions and their $t-$statistics thus cluster.

Second, under idea origination none of the unconditional alphas are significant. The maximal $t-$statistic in the lower panel in Figure 2 is well below the $10\%-$level. Therefore, on average the empiricist fails to identify abnormal returns when ideas are originated—significant performance (negative or positive) is virtually nonexistent.[14] Instead, idea sharing can explain empirical findings that a handful of managers generate significant alpha at the expense of a concentrated mass of significant underperformers (e.g., Barras et al. (2010)).

[14]This result is not specific to this calibration—it prevails under a vast range of parameters.

A broader issue is to understand under which conditions an empiricist fails to reject the null hypothesis of zero alpha for a perfect market timer; this involves determining the maximal level of statistical significance that prevails in a rational-expectations model. Maximal statistical significance in the model is determined by the $t-$statistic that a perfectly informed manager generates. This $t-$statistic is finite, although its associated alpha is not:

$$\widehat{t}_{max,t} \equiv \lim_{n\to\infty} \widehat{t}_{\alpha,t}(n) = \sqrt{\frac{2\mathbb{E}[\mathrm{SR}_t^2]}{\pi}} = \sqrt{\frac{2}{\pi}} \underbrace{|k_t|}_{\substack{\text{information-revelation}\\ \text{speed through prices}}} \underbrace{\sqrt{o_t + o_t^2(\phi_t/\sigma_S^2 + \gamma^2\sigma_\Theta^2(t+1))}}_{\text{scaled market-timing gains}}. \tag{31}$$

The proposition below characterizes the initial level as well as the asymptotic average level of this statistic in each model of idea gathering.

Proposition 6 (***Maximal level of statistical significance***)**.** *The maximal level of statistical significance in Eq.* (31) *at time* $t = 0$ *satisfies:*

$$\widehat{t}_{max,0} = \eta\sqrt{\frac{2}{\pi}}\frac{\sigma_\Pi\sqrt{\sigma_\Pi^2\left(\gamma^2\sigma_\Theta^2\sigma_S^2+1\right)^2+\gamma^2\sigma_\Theta^2\sigma_S^4}}{\sigma_\Pi^2+\gamma^2\sigma_\Theta^2\sigma_S^2\left(\sigma_\Pi^2+\sigma_S^2\right)}, \tag{32}$$

both under idea sharing and origination. Asymptotically, however, the average level of maximal statistical significance differs. Under idea origination it vanishes:

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T \widehat{t}_{max,s}\,ds = 0, \tag{33}$$

whereas, under idea sharing, statistical signifiance persists asymptotically:

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T \widehat{t}_{max,s}\,ds = 2\sqrt{\eta/\pi}. \tag{34}$$

Proof. Eq. (32) follows from substituting the expressions for the cross-sectional average ϕ in Corollaries 1 and 2 in Eq. (31). For Eqs. (33) and (34), refer to Corollary 7. □

Statistical significance improves with the intensity at which managers collect ideas. In particular, the initial level of statistical significance in Eq. (32) scales with the intensity η at which managers collect ideas in each model. Intuitively, as the information flow intensifies, stock return volatility decreases and statistical significance thus improves.[15]

Under idea origination, however, this effect is weak and temporary. When ideas are originated the resulting increase in knowledge does not feed back in the knowledge managers produce independently—the flow of private information is linear (see Corollary 1). As a result, the increase in η must be substantial to raise the maximal level of statistical significance in meaningful ways. Under a wide range of parameters, the information flow is not sufficiently intense to produce statistically significant performance. Most importantly, statistical significance vanishes asymptotically, as the average $t-$statistic in Eq. (33) indicates. Statistical significance is the product of the speed at which prices reveal information and (scaled) market-timing gains (see Eq. (31)). When the information flow is linear, information flows from prices at steady speed, while market-timing gains vanish asymptotically.

Unlike idea origination, idea sharing produces significant performance. Idea sharing allows knowledge to feed upon itself, which amplifies statistical significance and has a lasting effect on its average level. When managers share ideas, the information flow is measured in exponential units of calendar time (see Corollary 1). Prices thus reveal information at exponential speed, which offsets the decline in market-timing gains and permanently affects the average level of statistical significance in Eq. (34).

A relevant question is whether there exists a general condition that ensures performance persists asymptotically in a rational-expectations framework. There is, based on the relative rate of increase of the information flow, ϕ_t, as Proposition 7 demonstrates.

Proposition 7 (***Asymptotic persistence***). *Let ϕ_t be a continuous nondecreasing function*

[15]More broadly, statistical significance improves with trading aggressiveness. Reducing risk aversion, γ, noise trading, σ_Θ, or the noise in private information, σ_S in (32) has the same effect.

with asymptotic relative rate of increase bounded away from zero:

$$\lim_{t\to\infty} \phi_t'/\phi_t > 0. \tag{35}$$

A perfect market timer then maintains her alpha asymptotically, $\lim_{T\to\infty} \frac{1}{T}\int_0^T \widehat{t}_{max,s} ds > 0$.

Proof. A sufficient condition for Eq. (31) to be strictly positive asymptotically is that $\lim_{t\to\infty}(\phi_t')^2 o_t > 0$, which holds under Eq. (35) since it implies $\lim_{t\to\infty}\phi_t' > 0$ and

$$\lim_{t\to\infty}\int_0^t \left(\frac{\phi_s'}{\phi_t'}\right)^2 \mathrm{d}s \leq \lim_{t\to\infty}\left(\frac{\phi_t - \phi_0}{\phi_t'}\right)^2 < \infty \tag{36}$$

□

As long as information keeps flowing in relative terms asymptotically, a perfect market timer keeps delivering significant alpha. As a result, any polynomial flow of information results in vanishing asymptotic performance. Because idea sharing implies an exponential flow of information, information accumulates at a constant relative rate and perfect market timing thus keeps delivering significant alpha asymptotically.

5.1.2. The separation of skill from luck

In practice, the empiricist observes a finite sample of data and thus cannot compute the unbiased alpha of Proposition 5. Rather, she can estimate time series of *conditional* alpha (and obtain a proxy for the unconditional alpha by averaging this time series):

$$\alpha_t^i = \frac{\sigma_S(|k_t| + \gamma\sigma_\Theta)o_t^c}{\gamma(\sigma_S^2 + o_t^c\phi_t)}\sqrt{n_t^i}\Big(\underbrace{R_t(n_t^i)SR_t^2}_{\text{returns on skill}} + \underbrace{SR_t\epsilon_t^i}_{\text{returns on luck}}\Big), \tag{37}$$

which contains both her returns on skill and luck. The empiricist's ability to separate the two depends on the distribution from which alphas are drawn, as emphasized by Kosowski, Timmermann, Wermers, and White (2006).[16] Skill skews the distribution of alphas in proportion to the skill-to-luck ratio, whereas luck creates a mean-preserving spread around it.

[16] Other studies include Fama and French (2010), Barras et al. (2010) and Ferson and Chen (2015).

Because skilled managers take larger positions the spread around their skill is wider, making conditional alpha inappropriate for statistical inference. The scale of a manager's position, however, does not directly affect the $t-$statistic of the conditional alpha in (22):

$$t^{i}_{\alpha,t} = \text{sign}\left(\widehat{\theta}^{i}_{t}\right) \frac{\sigma^2_S |k_t|}{\sigma^2_S + o^c_t \phi_t} \left(\Delta_t + \gamma o^c_t \Theta_t\right) \equiv \text{sign}\left(\widehat{\theta}^{i}_{t}\right) \text{SR}_t, \tag{38}$$

which only depends on the direction of a manager's position, as opposed to its scale.[17] Importantly, since the direction of the average manager's position, $\text{sign}\left(\widehat{\theta}^{i}_{t}\right) = \text{sign}\left(\epsilon^{i}_{t}\right)$, follows a coin toss, her $t-$statistic defines the null hypothesis of pure luck.

Separating skill from luck in turn involves comparing the cross-sectional distribution of $t-$statistics that managers generate with that of the average manager, the theoretical counterpart to the bootstrap approach of Kosowski et al. (2006). Figure 3 makes this comparison under idea sharing (upper panel) and idea origination (lower panel), for the cross-sectional distribution of $t-$statistics averaged over the trading period. Each distribution is compared to that under the null of pure luck in each model (the red dashed line).

[insert Figure 3 here]

The separation of skill from luck concentrates in underperforming funds when ideas are shared, whereas skill is virtually indistinguishable from luck when ideas are originated. The distribution of cross-sectional $t-$statistics and the distribution under the null of pure luck differ in the upper panel only. Luck (dashed line) generates fewer negative $t-$statistics and more positive $t-$statistics—when ideas are shared the empiricist concludes that a manager is lucky when she outperforms and unskilled when she underperforms. In contrast, when ideas are originated (lower panel) the population of managers performs no better than luck.

Hence, unlike idea origination, idea sharing can explain why most underperforming funds appear to be truly unskilled, while most outperforming funds appear to be lucky (Kosowski

[17] Kosowski et al. (2006) make a related point, describing empirical $t-$statistics as pivotal statistics.

et al. (2006) and Fama and French (2010)).[18] To formalize this result, I examine the cross-sectional distribution of $t-$statistics at time t. The distributions in Figure 3 pertain to $t-$statistics averaged over the trading period and thus rely on simulations. In contrast, the distribution of $t-$statistics at a given time t can be characterized explicitly.

Proposition 8. *The cross-sectional distribution of conditional $t-$statistics at time t is a mixture of skew normal distributions weighted by the distribution of number of ideas:*

$$\mathbb{P}\left[t^i_{\alpha,t} \in (x, x+dx)\right] = l_t(x)\Big(1 + \sum_{k\in\mathbb{N}} c(k)(-1)^k x^{2k+1} \underbrace{\sum_{n\in\mathbb{N}^*} \mu_t(n)(R_t(n))^{2k+1}}_{\substack{\text{odd moments of the} \\ \text{skill-to-luck ratio}}}\Big), \tag{39}$$

for any $x \in \mathbb{R}$, where $c(\cdot)$ are positive, decreasing coefficients reported in Appendix G and $l_t(\cdot)$ is the p.d.f. of $t-$statistics under the null of pure luck; it is Gaussian and given by:

$$l_t(x) = \mathbb{P}\left[t^i_{\alpha,t} \in (x, x+dx)|n^i_t = \phi_t\right] = \frac{1}{\sqrt{2\pi\mathbb{E}[SR^2_t]}} \exp\left(-\frac{1}{2}\left(\frac{x}{\sqrt{\mathbb{E}[SR^2_t]}}\right)^2\right) dx. \tag{40}$$

As anticipated above when describing conditional alphas in Eq. (37), skill skews the distribution of $t-$statistics in the direction of the skill-to-luck ratio, while luck creates a mean-preserving spread around it. The distribution under the null of pure luck in Eq. (40) is Gaussian, with mean zero and variance equal to the expected, squared Sharpe ratio. The average manager flips a coin and pockets or loses the Sharpe ratio half of the time. The second term in Eq. (39) shifts mass from one side of the cross-sectional distribution to the other. The odd moments of the skill-to-luck ratio determines the direction of the resulting tilt—skill tilts the distribution to the left, lack thereof tilts it to the left.

How skewed the distribution is depends on the concavity of the skill-to-luck ratio and the shape of distribution of number of ideas. Under idea origination the cross-sectional number of ideas follows a Poisson distribution (see Corollary 2). It follows that for reasonably large

[18]The distribution of $t-$statistics in this model under idea sharing is strikingly similar to empirical findings. E.g., compare the upper panel of Figure 3 to the upper panel of Figure 2 in Kosowski et al. (2006).

intensities, η, this distribution is approximatively symmetric around the average manager. All odd moments in Eq. (39) are thus approximatively zero—the cross-sectional distribution of $t-$statistics and that under luck are indistinguishable.

Under idea sharing, however, the distribution of number of ideas (see Corollary 1) is strongly left-skewed. To see how this skew affects the distribution of $t-$statistics, focus on $t-$statistics that are close to zero, for which the first moment in Eq. (39) dominates:

$$\mathbb{P}\left[t^i_{\alpha,t} \in (x, x+\mathrm{d}x)\right] = l_t(x)\big(1 + \sqrt{2/\pi}\ \underbrace{\mathbb{E}[R_t(n)]}_{\substack{\leq 0 \text{ by Jensen's} \\ \text{inequality}}}\ x\big) + O(x^3). \tag{41}$$

The concavity of the skill-to-luck ratio and Jensen's inequality jointly imply that the first moment of this ratio is negative. That the distribution $\mu(\cdot)$ is strongly skewed further implies that this moment is large. As a result, the second term in Eq. (41) shifts mass from the right to the left of the cross-sectional distribution of $t-$statistics, tilting it to the left.

Another way of understanding this result—skill separates from luck for underperforming funds only—is to consider the probability that a manager who has n ideas times the market successfully, i.e., the probability that her $t-$statistic is positive (see Appendix H):

$$\mathbb{P}\left[t^i_{\alpha,t} > 0 \middle| n^i_t = n\right] = \frac{1}{2} + \frac{1}{\pi}\tan^{-1}\left(\sqrt{\mathbb{E}\left[SR^2_t\right]}R_t(n)\right). \tag{42}$$

The probability in Eq. (42) shows that a manager's success at timing the market depends on the sign of her skill-to-luck ratio. Because the average manager's skill-to-luck ratio is zero, her strategy is no better than a coin toss—she is on the wrong side of the market half of the time. It follows that unskilled managers, whose skill-to-luck ratio is negative, do strictly worse at timing the market than a coin toss and skilled managers do strictly better.

[insert Figure 4 here]

How much better or worse a strategy does relative to a coin toss depends on the concavity of the skill-to-luck ratio and the size of the squared Sharpe ratio. In Figure 4 I plot the probability in (42) as a function of a manager's number of ideas. For the illustration I denote by $\overline{SR} \geq \mathbb{E}[SR_t^2]$ the largest market-timing gains that can be reaped in this model.

Suppose first that the squared Sharpe ratio is larger than $\overline{SR}$. The probability of success then polarizes around the average manager and skill cleanly separates from luck, as the blue line in Figure 4 shows. However, this situation cannot be an equilibrium outcome: unskilled managers would refuse to trade, knowing that skilled managers, who all are perfect market timers, would always take advantage of them. The squared Sharpe ratio must be thus inferior to $\overline{SR}$ for markets to clear, in which case the probability of success becomes markedly concave, as the red line indicates. An additional idea improves the odds of success for unskilled managers by significantly more than it does for skilled managers.

An increase in the information flow, ϕ_t, impairs the separation of skill from luck in two ways; it simultaneously reduces the squared Sharpe ratio and increases the concavity of the skill-to-luck ratio. Under idea sharing, both effects increase the concavity of the probability of success substantially, causing the separation of skill from luck to concentrate in underperforming funds. Proposition 9 further shows that the mass of $t-$statistics is always concentrated on the left side of the cross-sectional distribution under idea sharing.

Proposition 9. *Let the distribution of number of ideas, $\mu_t(\cdot)$, satisfy Corollary 1 (idea sharing) with $\eta > 0$. At any finite time the mass of the cross-sectional distribution of $t-$statistics is shifted to the left relative to the null hypothesis of pure luck:*

$$\mathbb{P}[t^i_{\alpha,t} \leq 0] > 1/2. \tag{43}$$

This result extends to all symmetric distributions of number of ideas.

Whether the result of Proposition 9 extends to alternative distributions of ideas depends on their shape. For instance, suppose this distribution has multiple modes; this typically occurs when managers form networks, leading ideas to cluster (e.g., the setup in Section 5.2.2). In this case, it is possible to shift mass from one mode to another so as to flip the inequality in Eq. (43), which may give the empiricist an opportunity to detect skill.

Whether such opportunities exist is an empirical debate. Using a same bootstrap approach, Kosowski et al. (2006) find strong evidence of skill, which materializes as a "shoulder" in the right tail, whereas Fama and French (2010) find weak evidence of skill in the extreme right tail. In this model, even in the presence of networks, skill remains hardly detectable.

5.2. Time-series implications and performance persistence

In this section I analyze how managers' performance depends on time. To reduce cross-sectional heterogeneity, I segment the population of managers into two groups, A and B. I then examine persistence when this segmentation is exogenous and when it arises endogenously through networks. The main insight is that network formation can explain why performance persistence concentrates in underperforming funds (Carhart, 1997).

The segmentation rule is as follows: a manager belongs to Group A or B depending on whether her number n of ideas is in $A = \{n \in \mathbb{N}^* : n < N\}$ or $B = \{n \in \mathbb{N}^* : n \geq N\}$, i.e., whether she holds more or fewer ideas than some threshold number, N. For the average manager of each group I then compute an alpha and its $t-$statistic (see Proposition 5):

$$\widehat{\alpha}_t^{\mathcal{K}} = \widehat{\alpha}_t \left(\mathbb{E}\left[n_t^i \middle| n_t^i \in \mathcal{K}\right]\right) \qquad \text{and} \qquad \widehat{t}_{\alpha,t}^{\mathcal{K}} = \widehat{t}_{\alpha,t} \left(\mathbb{E}\left[n_t^i \middle| n_t^i \in \mathcal{K}\right]\right), \qquad \mathcal{K} = A, B. \tag{44}$$

An important determinant of how these statistics evolve over time is how the population is segmented; whether ideas are shared or originated matters asymptotically for the maximal $t-$statistic (see Proposition 6), but leads to qualitatively similar patterns over finite periods of time across the two groups. I thus focus on idea sharing exclusively to contrast the evolution of alpha and its $t-$statistic under exogenous and endogenous segmentation.

5.2.1. *Exogenous Segmentation*

I start by splitting the population exogenously into skilled and unskilled managers. That is, based on Definition 2 the segmentation threshold N is the cross-sectional average ϕ; unskilled managers belong to Group A and skilled managers belong to Group B. In Figure 7, I plot alpha and its $t-$statistic in Eq. (44) as a function of time, both for the average manager in Group A (the dashed red line) and B (the solid black line).

[insert Figure 5 here]

Under exogenous segmentation, performance converges across groups of skilled and unskilled managers. Eq. (23) indicates that a manager maintains her alpha only if she offsets the decline in the squared Sharpe ratio with a sufficient improvement in her skill. Over time skill deteriorates in Group A and improves in Group B. However, the decline in the squared Sharpe ratio dominates the improvement or deterioration in skill in each group. As a result, alphas become gradually indistinguishable across groups (left panel of Figure 5).

Importantly, alphas and their $t-$statistics move symmetrically across skilled and unskilled managers. This symmetry is perhaps surprising in light of the concavity of the skill-to-luck ratio—a decline or an improvement in skill affects $t-$statistics comparitively more for unskilled managers than it does for skilled managers. However, skill in Group A does not deteriorate as fast as it improves in Group B, which offsets the asymmetry implied by the skill-to-luck ratio. It follows that performance, be it negative or positive, does not persist under exogenous segmentation. I now show how network formation affects this outcome.

5.2.2. *Endogenous Segmentation*

For the moment I have assumed that all managers belong to a same network and have the same ability to gather ideas. However, evidence indicates that fund managers interact within networks. For instance, managers form investor “cliques” (Crane, Koch, and Michenaud, 2015).[19] To introduce networks in the model I follow Duffie, Malamud, and Manso (2015) and assume that the population of managers is segmented into different classes that differ by quality of information and "connectivity". As in the previous section, I consider only two classes of managers, which I call Network A and B. In this section, however, the segmentation rule determines how managers move across networks and how each network arises.

I interpret the segmentation rule as "valuable investment ideas" remaining localized among a small group of managers (Stein, 2008). In this CARA-normal setup, "valuable" is a synonym for "precise". I define a precise idea as an average idea composed of N or more ideas. I do not model incentives for managers to keep precise ideas localized in a small group.[20] The segmentation rule then implies that precise ideas are located in Network B. Under the broader interpretation that managers buy the same research, segmentation means that some institutions, say larger ones, receive more reports and end up having better ideas.

The defining feature of a network is that managers within a same network are more likely to be matched—they are more connected. For instance, think about each network as a mutual fund family. Members are mostly matched within their own family, but sometimes are matched to members of other families at the annual CFA meeting. Consistent with this idea, pairs of managers are sampled from the same network with probability $p \in (1/2, 1]$ and from different networks with probability $1 - p$. The network affiliation of each manager involved in a pair is common knowledge (Duffie et al., 2015).

A manager is randomly matched with an intensity that is proportional to her number of potential matches. This assumption introduces differential meeting intensities within and outside networks in the model. Formally, let $q_t \in [0, 1]$ be the (endogenously determined) fraction of the population that is located in Network A at time t. Let the parameter, $\eta/2$, represent the matching intensity in both networks when of identical size. A manager with n

[19] Examples include educational network (Cohen et al., 2008), interactions within the same neighborhood (Pool et al., 2015), the same city (Hong et al., 2005; Christoffersen and Sarkissian, 2009), or within investment forums, e.g.,The Value Investor Club, The Alburn Village or The Q Group (Gray et al., 2012).

[20] I could add to the model a utility cost for sharing a large number of ideas, but I have avoided this for tractability. Endogenizing the concept of "precise idea" would also likely make the threshold N time-varying and manager-specific. For tractability I take the threshold N to be fixed and common to all managers.

ideas is then matched to someone within her own network with intensity $\eta p(\mathbf{1}_{n\in A}q_t + \mathbf{1}_{n\in B}(1-q_t))$ and to someone outside her network with intensity $\eta(1-p)(\mathbf{1}_{n\in A}(1-q_t)+\mathbf{1}_{n\in B}q_t)$. Figure 6 illustrates the resulting matching process and network structure.

[insert Figure 6 here]

Based on this structure the information flow in Eq. (4) takes the following form.

Corollary 4. ***(Network formation)*** *Under idea sharing with network formation, the cross-sectional average number ϕ of ideas satisfies*

$$\phi_t = \frac{e^{\eta t}}{b_t}\left(1+\eta(2p-1)\int_0^t e^{-\eta s}\left(2b_s^{-1}\frac{a_s^{N-1}-1}{a_s-1}-1\right)\frac{1+a_s^{N-1}(a_s(N-1)-N)}{(1-a_s)^2}ds\right) \quad (45)$$

where a_t and b_t are deterministic coefficients, the explicit solution of which is in the appendix. The solution for the distribution μ of ideas is also relegated to the appendix.

An expected consequence of information segmentation between Network A and B is that it dampens the percolation of ideas. Absent information segmentation, the cross-sectional average, ϕ, increases exponentially at rate η (see Corollary 27). Whereas this exponential increase also appears in Eq. (45), it is multiplied by a factor that slows percolation down. To illustrate how dampening affects performance persistence, I plot the counterpart to Figure 5 with endogenous segmentation in Figure 7.

[insert Figure 7 here]

Alphas converge across networks, whereas their $t-$statistics do not. Furthermore, a manager's alpha and its $t-$statistic may move in opposite directions, which they do in Network A but not in Network B. Under network formation, skill deteriorates at an accelerated rate in Network A and improves at a decaying rate in Network B. As a result, the asymmetry implied by the concavity of the skill-to-luck ratio now dominates; $t-$statistics become strongly negative in Network A and converge to zero in Network B (right panel). Performance persistence thus concentrates in underperforming funds (e.g., Carhart (1997)).

These results also show that $t-$statistics may or may not have superior time-series properties relative to alphas depending on how managers gather ideas. When managers form networks, $t-$statistics retain the ability to rank them consistently even when their alphas converge, whereas $t-$statistics do not under exogenous segmentation.

6. Extensions: fund flows and fees

In this section I incorporate two important institutional features of mutual funds—fund flows and fees. Because mutual funds' performance fees are of the fulcrum type, they do not affect the results in a CARA-normal framework. Whereas fund flows cause managers to herd towards their benchmark, they leave the results qualitatively unchanged.

To model fund flows I adopt the reduced-form approach in Koijen (2012). Flows, $F(W_T^i, B_T)$, in and out of fund i are linear in performance relative to some benchmark, B. In this CARA-normal setup, I work with dollar performance, as opposed to returns:

$$F(W_T^i, B_T) = \tau(W_T^i - B_T), \tag{46}$$

where $\tau \geq 0$ is a flow-performance parameter. That is, funds flow in or out as fund i out- or under-performs the benchmark, respectively. Note that I could add a fixed part to the flow function in Eq. (46); however, it would not affect equilibrium outcomes in this framework. It follows that total assets under management (AUM) at the horizon date equal:

$$\text{AUM}_T^i = W_T^i + F(W_T^i, B_T). \tag{47}$$

I assume that each fund i charges fees, $f \times \text{AUM}_T^i$, in proportion to assets under management, taking the fee rate, f, as exogenous (e.g., Hugonnier and Kaniel (2010) and Koijen

(2012)). This fee specification is known as *fulcrum* performance fees—bonuses for outperforming the benchmark are symmetric to the penalties for underperforming it. As Cuoco and Kaniel (2011) point out, performance fees in the mutual industry are required to be of the fulcrum type. I assume that managers maximize expected utility over their compensation:

$$\mathbb{E}\left[-\exp\left(-\gamma f(W_T^i + F(W_T^i, B_T))\right)\middle|\, \mathscr{F}_t^i\right]. \tag{48}$$

Under this specification, fees act as a scale on managers' risk aversion.

I view the benchmark, B, as individual investors' effort to benchmark managers against each other. Absent information asymmetry, individual investors could benchmark each fund against average managed wealth at the horizon date T:

$$\overline{W}_T = \int_0^1 W_T^i \mathrm{d}i = \int_0^T \Theta_s \mathrm{d}P_s + \Theta_T \Delta P_T. \tag{49}$$

Information asymmetry makes this comparison impossible—managers do not observe average managed wealth, nor do individual investors. However, individual investors can form a proxy for average managed wealth based on publicly available information $\mathscr{F}^c$:

$$B_T \equiv \int_0^T \widehat{\Theta}_s^c \mathrm{d}P_s + \widehat{\Theta}_{T-}^c \Delta P_T, \tag{50}$$

a benchmark that any investor can use in real time. Under this specification, fund flows act as a simple form of relative wealth concerns (Garcia and Strobl, 2011).[21]

I now sketch how these two features affect the solution to the baseline model in Section 3, relegating computational details to Appendix I. While flows do not directly modify the

[21] In a static rational-expectations setup Garcia and Strobl (2011) model wealth concerns relative to average managed wealth, as opposed to the commonly available proxy in Eq. (50).

learning dynamics in Proposition 1, they modify managers' trading strategy:

$$\theta_t^i \equiv \frac{1}{f(\tau+1)}\theta_t^{i,\text{baseline}} + \frac{\tau}{\tau+1}\widehat{\Theta}_t^c, \tag{51}$$

where $\theta^{i,\text{baseline}}$ is the portfolio in Proposition 2. Strategies now incorporate an additional component, which tilts managers' position towards the proxy for the average trade (the benchmark); the flow-performance parameter determines the strength of the tilt. Intuitively, managers suffer from underperforming their peers due to the resulting outflows. They hedge against outflows by herding towards what individual investors perceive as the average strategy. Furthermore, fees scale the baseline position by adjusting risk aversion.

Clearing markets the modified portfolio in Eq. (51) in turn modifies price coefficients:

$$\lambda_{1,t} = \frac{1}{\tau+1}\lambda_{1,t}^{\text{baseline}} \quad \text{and} \quad \lambda_{2,t} = f\lambda_{2,t}^{\text{baseline}}, \tag{52}$$

in which the expression for common uncertainty is also modified (see Appendix I). Because fund flows cause managers to herd towards the benchmark, they reduce price informativeness. The first price coefficient, λ_1, scales down relative to the baseline case. By adjusting risk aversion fees further scale the risk premium through the second price coefficients, λ_2. The signal-noise ratio, λ_1/λ_2 may increase or decrease depending on which effect dominates.

The presence of fund flows breaks the proportionality between managers' trading intensity on information and the Sharpe ratio. Substituting the price coefficients in Eq. (52) into the portfolio in Eq. (51) and deriving informational holdings yields (see Appendix I):

$$\widehat{\theta}_t^i = \frac{1}{f\gamma(\tau+1)}\left(\frac{n_t^i-\phi_t}{\sigma_S^2+o_t^c\phi_t}\left(\frac{\sigma_S^2(\tau+1)+\tau o_t^c\phi_t}{\sigma_S^2(\tau+1)}\Delta_t + f\gamma o_t^c\Theta_t\right) + \frac{\sqrt{n_t^i}}{\sigma_S}\epsilon_t^i\right) + \tau H_t. \tag{53}$$

The term in bracket is similar to the expression in Proposition 4 and shows that the concept of skill in Definition 2 remains valid in the presence of flows. However, herding in response to outflows creates an additional demand, H, which is common to all managers and contributes to break the proportionality of informational demands to the Sharpe ratio.

Due to the resulting loss of tractability, I proceed with numerical illustrations onwards. Since fulcrum fees simply scale risk aversion in this model, they are ignored in numerical illustrations. Note, however, that relaxing the symmetry between bonuses and penalties may affect the results in nontrivial ways. For instance, suppose managers receive compensation for outperforming their peers, but do not incur penalties for underperforming them. Based on the intuition in Cuoco and Kaniel (2011), managers have incentives to increase tracking error when their performance approaches the benchmark. This risk-shifting behavior would presumably make performance noisier at the center of the distribution—it is unlikely to affect the asymmetry between skilled and unskilled managers.

To illustrate the effect of fund flows I focus on idea sharing. Furthermore, I use the estimate in Koijen (2012) for the performance-flow parameter, $\tau = 0.86$. This estimate implies that for each dollar a fund outperforms the benchmark, it attracts 86 cents in flows. I repeat the steps of Section 5.1.2 in the presence of fund flows and plot the result in Figure 8.

[insert Figure 8 here]

Fund flows do not affect cross-sectional implications qualitatively. Relative to Figure 3, the left panel of Figure 8 is virtually unchanged. Intuitively, flows induce funds to herd through the second term in Eq. (51). This term does not depend on a manager's number of ideas. As a result, the portfolio tilt towards the benchmark is identical for skilled and unskilled managers alike and keeps cross-sectional predictions qualitatively unchanged.

By the same token, the time evolution of alphas and their $t-$statistics is similar in the presence of fund flows. Just like in Figure 5 alphas and $t-$statistics converge. Only the magnitude and rate of convergence differ. This difference is primarily due to the effect of flows on price informativeness. Herding towards the benchmark reduces both the speed at which prices reveal information and the level of price informativeness. As a result, flows slow down the convergence of alpha and its $t-$statistic and reduce their magnitude.

7. Conclusion

This paper shows that idea sharing, as opposed to idea origination, can explain stylized facts regarding fund managers' performance—the separation of skill from luck, statistical significance of alpha and its persistence concentrate in the worst-performing funds.

Methodologically, this paper develops a new framework that embeds a mechanism of discrete information collection in a standard continuous time, rational-expectations equilibrium model. That private information arrives discretely in a framework in which trading occurs continuously produces a novel form of learning—public information continuously flows from prices, while private information flows at discrete, random times. I prove that a rational-expectations framework extends to this broader class of continuous-discrete filtering processes, while keeping its tractability (the equilibrium is solved in closed form).

I hope that this framework can be used to explore other equilibrium consequences of idea sharing. For instance, it may be adapted for the purpose of studying how rumors propagate, driving prices away from fundamental values. Allowing agents to act strategically would lead to a new theory of price manipulation. Moreover, the relation between idea sharing and prices is only one-way—idea sharing impacts prices, but prices do not impact the way in which agents interact. A challenging extension involves a full-fledged equilibrium in which idea sharing and prices feed back both ways. Idea sharing may also mitigate "limits to arbitrage"—it may help arbitrageurs to synchronize their trades, reducing convergence risk.

Appendix A. Proof of Proposition 1 (Learning)

In this appendix, I derive the dynamics of the Bayesian updating process for an agent who obtains private information according to the mechanism described in Section 2.1. I show that the Bayesian updating process features both continuous updates and discrete updates occurring at random times.

Let $X_t = (\Pi, \Theta_t)^\top$ denote the unobservable state vector with dynamics

$$\begin{aligned} \mathrm{d}X_t &= \begin{bmatrix} 0 & 0 \\ 0 & -a_\Theta \end{bmatrix} X_t \mathrm{d}t + \begin{bmatrix} 0 \\ \sigma_\Theta \end{bmatrix} \mathrm{d}B_t^\Theta \\ &\equiv aX_t \mathrm{d}t + b\mathrm{d}B_t^\Theta. \end{aligned}$$

Agents observe two types of signal. First, they receive continuous news updates through the price P_t. Using (9), define the normalized price signal ξ_t as

$$\xi_t \equiv P_t - (1-\lambda_{1,t})\widehat{\Pi}_t^c = \lambda_{1,t}\Pi + \lambda_{2,t}\Theta_t$$

and notice that $\sigma(P_s : 0 \le s \le t) \Leftrightarrow \sigma(\xi_s : 0 \le s \le t)$. Applying Ito's lemma, the dynamics of ξ_t satisfy

$$\begin{aligned} \mathrm{d}\xi_t &= \begin{bmatrix} \lambda'_{1,t} & \lambda'_{2,t} - a_\Theta \lambda_{2,t} \end{bmatrix} X_t \mathrm{d}t + \lambda_{2,t}\sigma_\Theta \mathrm{d}B_t^\Theta \\ &\equiv A_{1,t} X_t \mathrm{d}t + B_{1,t}\mathrm{d}B_t^\Theta. \end{aligned}$$

Second, each agent i receives a sequence of private signals at Poisson arrival times (i.e., the times at which the Poisson counter N^i increases):

$$\{0 = \tau_0 \le \tau_1 \le \tau_2 \le ... \le \tau_{N_T^i} \le \tau_{N_T^i+1} = T\}$$

and gradually builds a collection $(S_j^i : 1 \le j \le n_t^i)$ of n_t^i discrete signals at time t. Agent i's information set $\mathscr{F}_t^i$ at time t is therefore given by

$$\mathscr{F}_t^i = \sigma\left(\left(\xi_s, S_j^i\right) : 0 \le s \le t, 1 \le j \le n_t^i\right), \ \ 0 \le t < T. \tag{54}$$

Accordingly, define the conditional mean $\widehat{X}_t^i = \mathbb{E}\left[X_t \middle| \mathscr{F}_t^i\right]$ and the positive semi-definite conditional variance-covariance matrix $O_t^i = \mathbb{E}\left[(X - \widehat{X}_t^i)(X - \widehat{X}_t^i)^\top \middle| \mathscr{F}_t^i\right]$ of the state vector X_t with respect to $\mathscr{F}_t^i$. At time $t \in (\tau_k, \tau_{k+1})$, for any $k \in \{0, 1, ..., N_T^i\}$, $\widehat{X}_t^i$ and O_t^i satisfy the dynamics given in Theorem A.1.

Theorem A.1. *The conditional mean* $\widehat{X}_t^i$ *with respect to the filtration* $\mathscr{F}_t^i$, $\forall k \in \{0, 1, ..., N_T^i\}$, *has dynamics*

$$d\widehat{X}_t^i = a\widehat{X}_t^i dt + (O_t^i A_{1,t}^\top + bB_{1,t}^\top)(B_{1,t}B_{1,t}^\top)^{-\frac{1}{2}} d\widehat{B}_t^i, \quad \forall t \in (\tau_k, \tau_{k+1}) \tag{55}$$

where the conditional variance-covariance matrix O_t^i *is given by the solution to the Ricatti equation*

$$\dot{O}_t^i = aO_t^i + O_t^i a^\top + bb^\top - (O_t^i A_{1,t}^\top + bB_{1,t}^\top)(B_{1,t}B_{1,t}^\top)^{-1}(A_{1,t}O_t^i + B_{1,t}^\top b), \quad \forall t \in (\tau_k, \tau_{k+1}) \tag{56}$$

and where the filter innovation $\widehat{B}_t^i$ satisfying

$$d\widehat{B}_t^i = (B_{1,t}B_{1,t}^\top)^{-\frac{1}{2}}(d\xi_t - A_{1,t}\widehat{X}_t^i dt) \tag{57}$$

is a $\widehat{\mathbb{P}}^i-$Brownian motion with respect to the filtration $\mathscr{F}_t^i$.

Proof. See Liptser and Shiryaev (2001), Theorem 12.7. □

Remark 1. *The dynamics in (55) and (56) can be simplified: since $\xi_t \in \mathscr{F}_t^i$, it follows that $\xi_t = \lambda_{1,t}\Pi + \lambda_{2,t}\Theta_t = \lambda_{1,t}\widehat{\Pi}_t^i + \lambda_{2,t}\widehat{\Theta}_t^i$ and the variance-covariance matrix O_t^i therefore satisfies*[22]

$$O_t^i = \begin{bmatrix} 1 & -\frac{\lambda_{1,t}}{\lambda_{2,t}} \\ -\frac{\lambda_{1,t}}{\lambda_{2,t}} & \left(\frac{\lambda_{1,t}}{\lambda_{2,t}}\right)^2 \end{bmatrix} o_t^i \equiv \Omega_t o_t^i \tag{58}$$

where $o_t^i = \mathbb{E}\left[(\Pi - \widehat{\Pi}_t^i)^2 \middle| \mathscr{F}_t^i\right]$. Using (58), the dynamics of the conditional mean in (55) simplify to

$$\begin{aligned} d\widehat{X}_t^i &= a\widehat{X}_t^i dt + \frac{1}{\lambda_{2,t}^2\sigma_\Theta}\begin{bmatrix} o_t^i(\lambda_{1,t}'\lambda_{2,t} - \lambda_{1,t}(\lambda_{2,t}' - a_\Theta\lambda_{2,t})) \\ \lambda_{2,t}^2\sigma_\Theta^2 + o_t^i\left(\frac{\lambda_{1,t}^2}{\lambda_{2,t}}(\lambda_{2,t}' - a_\Theta\lambda_{2,t}) - \lambda_{1,t}\lambda_{1,t}'\right)\end{bmatrix} d\widehat{B}_t^i & (59) \\ &= a\widehat{X}_t^i dt + \begin{bmatrix} o_t^i k_t \\ \left(\sigma_\Theta - o_t^i\frac{\lambda_{1,t}}{\lambda_{2,t}}k_t\right)\end{bmatrix} d\widehat{B}_t^i, \quad \forall t \in (\tau_l, \tau_{l+1}),\ l \in \{0, 1, ..., N_T^i\} & (60) \end{aligned}$$

where

$$k_t \equiv \frac{1}{\sigma_\Theta}\left(\frac{d}{dt}\left(\frac{\lambda_{1,t}}{\lambda_{2,t}}\right) + a_\Theta\frac{\lambda_{1,t}}{\lambda_{2,t}}\right). \tag{61}$$

Similarly, the matrix Riccati equation for the variance-covariance matrix O_t^i in (56) simplifies to an ordinary Riccati equation for o_t^i:

$$\frac{do_t^i}{dt} = -\frac{(o_t^i)^2}{\lambda_{2,t}^4\sigma_\Theta^2}(\lambda_{1,t}'\lambda_{2,t} - \lambda_{1,t}(\lambda_{2,t}' - a_\Theta\lambda_{2,t}))^2 = -k_t^2(o_t^i)^2, \quad \forall t \in (\tau_l, \tau_{l+1}),\ l \in \{0, 1, ..., N_T^i\}. \tag{62}$$

At time $t = \tau_k$, $\forall k \in \{0, 1, ..., N_T^i\}$, agent i receives a sequence $(S_{j+n_{t-}^i}^i : 1 \le j \le \Delta n_t^i)$ of Δn_t^i new signals and her information set jumps to

$$\mathscr{F}_t^i = \mathscr{F}_{t-}^i \vee \sigma((S_{j+n_{t-}^i}^i : 1 \le j \le \Delta n_t^i)),$$

[22] Notice that $\Theta_t - \widehat{\Theta}_t^i = -\frac{\lambda_{1,t}}{\lambda_{2,t}}(\Pi - \widehat{\Pi}_t^i)$ implies that

$$\mathbb{E}\left[(\Theta_t - \widehat{\Theta}_t^i)^2 \middle| \mathscr{F}_t^i\right] = \mathbb{E}\left[\left(\frac{\lambda_{1,t}}{\lambda_{2,t}}\right)^2(\Pi - \widehat{\Pi}_t^i)^2 \middle| \mathscr{F}_t^i\right] = \left(\frac{\lambda_{1,t}}{\lambda_{2,t}}\right)^2 o_t^i$$

and

$$\mathbb{E}\left[(\Theta_t - \widehat{\Theta}_t^i)(\Pi - \widehat{\Pi}_t^i) \middle| \mathscr{F}_t^i\right] = \mathbb{E}\left[-\frac{\lambda_{1,t}}{\lambda_{2,t}}(\Pi - \widehat{\Pi}_t^i)^2 \middle| \mathscr{F}_t^i\right] = -\frac{\lambda_{1,t}}{\lambda_{2,t}}o_t^i.$$

where $\mathscr{F}^i_{t-} = \sigma\left(\bigcup_{s<t} \mathscr{F}^i_s\right)$. That is, agent i's filtration is not left-continuous $\mathscr{F}^i_t \neq \mathscr{F}^i_{t-}$, i.e., there is a "surprise" in agent i's information flow. This surprise is the sequence of new signals, which can be expressed as an aggregate signal according to Lemma A.1.

Lemma A.1. *Conditional on $t = \tau_k$, $\forall k \in \{0, 1, ..., N^i_T\}$, define agent i's aggregate signal Y^i_t as*

$$Y^i_t = (\Delta n^i_t)^{-1} \sum_{j=1}^{\Delta n^i_t} S^i_{j+n^i_{t-}} = \left[\begin{array}{cc} 1 & 0 \end{array}\right] X_t + \frac{\sigma_S}{\sqrt{\Delta n^i_t}} \epsilon^i_t \tag{63}$$

$$\equiv A_{2,t} X_t + B_{2,t}(\Delta n^i_t) \epsilon^i_t \tag{64}$$

where $\Delta n^i_t \sim \pi_t(\cdot; n^i_{t-})$, $\epsilon^i_t \sim \mathscr{N}(0,1)$ and $\epsilon^i_{\tau_k} \perp \epsilon^i_{\tau_l}$, $\forall k \neq l$. Then, the aggregate signal Y^i_t is a sufficient statistic for the sequence $(S^i_{j+n^i_{t-}} : 1 \leq j \leq \Delta n^i_t)$.

Proof. Denote by $p(\Pi|\mathscr{F}^i_t)$ the conditional density of Π with respect to $\mathscr{F}^i_t$. Fixing a time $t = \tau_k$ and applying Bayes' rule, the conditional density $p(\Pi|\mathscr{F}^i_t)$ satisfies the recursive relation

$$p(\Pi|\mathscr{F}^i_t) = \frac{p(\Pi|\mathscr{F}^i_{t-}) f(\mathbf{S}|\Pi)}{\int_{\mathbb{R}} p(x|\mathscr{F}^i_{t-}) f(\mathbf{S}|x) \mathrm{d}x} \tag{65}$$

where $f(\mathbf{S}|\Pi)$ denotes the density of a vector of signals $\mathbf{S}$ conditional on Π and where $p(\Pi|\mathscr{F}^i_{t-})$ satisfies

$$p(\Pi|\mathscr{F}^i_{t-}) = \left(2\pi o^i_{t-}\right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \frac{(\widehat{\Pi}^i_{t-} - \Pi)^2}{o^i_{t-}}\right) \tag{66}$$

since, from Theorem A.1, $p(\Pi|\mathscr{F}^i_t)$ is conditionally Gaussian for any $t \in (\tau_{k-1}, \tau_k)$ (Liptser and Shiryaev (2001), Theorem 12.6). First, let $\mathbf{S} = \left[\begin{array}{cccc} S^i_{n^i_{t-}+1} & S^i_{n^i_{t-}+2} & \cdots & S^i_{n^i_{t-}+\Delta n^i_t} \end{array}\right]^\top$ be the vector of signals in the sequence $(S^i_{j+n^i_{t-}} : 1 \leq j \leq \Delta n^i_t)$. Conditional on Π, these signals are independent and thus

$$f(\mathbf{S}|\Pi) = (2\pi\sigma_S^2)^{-\frac{\Delta n^i_t}{2}} \prod_{j=1}^{\Delta n^i_t} \exp\left(-\frac{1}{2}\left(\frac{S^i_{j+n^i_{t-}} - \Pi}{\sigma_S}\right)^2\right). \tag{67}$$

After substituting (67) in (65) and integrating, the conditional density $p(\Pi|\mathscr{F}^i_t)$ is explicitly given by

$$p(\Pi|\mathscr{F}^i_t) = \sqrt{\frac{1}{2\pi}\left(\frac{1}{o^i_{t-}} + \frac{\Delta n^i_t}{\sigma_S^2}\right)} \exp\left(-\frac{1}{2}\left(\frac{1}{o^i_{t-}} + \frac{\Delta n^i_t}{\sigma_S^2}\right)\left(\left(\frac{\widehat{\Pi}^i_{t-}}{o^i_{t-}} + \frac{\sum_{j=1}^{\Delta n^i_t} S^i_{j+n^i_{t-}}}{\sigma_S^2}\right)\left(\frac{1}{o^i_{t-}} + \frac{\Delta n^i_t}{\sigma_S^2}\right)^{-1} - \Pi\right)^2\right). \tag{68}$$

Second, let $\mathbf{S} = Y^i_t$ be the aggregate signal, in which case

$$f(Y^i_t|\Pi) = \left(2\pi\frac{\sigma_S^2}{\Delta n^i_t}\right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\left(Y^i_t - \Pi\right)^2 \frac{\Delta n^i_t}{\sigma_S^2}\right). \tag{69}$$

Substituting (69) in (65) and integrating, the conditional density in (68) becomes

$$p(\Pi|\mathscr{F}_t^i) = \sqrt{\frac{1}{2\pi}\left(\frac{1}{o_{t-}^i} + \frac{\Delta n_t^i}{\sigma_S^2}\right)} \exp\left(-\frac{1}{2}\left(\frac{1}{o_{t-}^i} + \frac{\Delta n_t^i}{\sigma_S^2}\right)\left(\left(\frac{\widehat{\Pi}_{t-}^i}{o_{t-}^i} + \frac{Y_t^i}{\sigma_S^2/\Delta n_t^i}\right)\left(\frac{1}{o_{t-}^i} + \frac{\Delta n_t^i}{\sigma_S^2}\right)^{-1} - \Pi\right)^2\right). \tag{70}$$

Clearly, the expressions in (68) and (70) are equivalent if and only if $Y_t^i = (\Delta n_t^i)^{-1} \sum_{j=1}^{\Delta n_t^i} S_{j+n_{t-}^i}^i$ and, by induction, this result must be true for all $t = \tau_k$ and $k \in \{0, 1, ..., N_T^i\}$. □

Using the result of Lemma A.1, at time $t = \tau_k$, for any $k \in \{0, 1, ..., N_T^i\}$, $\widehat{X}_t^i$ and O_t^i are updated according to Theorem A.2.

Theorem A.2. *The conditional mean $\widehat{X}_t^i$ with respect to the filtration $\mathscr{F}_t^i$, $\forall k \in \{0, 1, ..., N_T^i\}$, satisfies*

$$\widehat{X}_t^i = \widehat{X}_{t-}^i + O_{t-}^i A_{2,t}^\top \left(A_{2,t} O_{t-}^i A_{2,t}^\top + B_{2,t}(\Delta n^i) B_{2,t}(\Delta n^i)^\top\right)^{-1} \widehat{Y}_t^i, \quad \forall t = \tau_k \tag{71}$$

where the conditional variance-covariance matrix O_t^i satisfies

$$O_t^i = O_{t-}^i - O_{t-}^i A_{2,t}^\top \left(A_{2,t} O_{t-}^i A_{2,t}^\top + B_{2,t}(\Delta n^i) B_{2,t}(\Delta n^i)^\top\right)^{-1} A_{2,t} O_{t-}^i, \quad \forall t = \tau_k \tag{72}$$

and where the filter innovation $\widehat{Y}_t^i$ satisfying

$$\widehat{Y}_t^i = \left(Y_t^i - E\left[Y_t^i | \mathscr{F}_{t-}^i, \Delta n_t^i\right]\right) \sim \mathscr{N}\left(0, o_{t-}^i + \frac{\sigma_S^2}{\Delta n_t^i}\right) \tag{73}$$

is normally distributed conditional on the filtration $\mathscr{F}_{t-}^i$ and Δn_t^i.

Proof. The updating rules in (71) and (72) directly follow from the proof of Lemma A.1: observe that (66) must also hold at time $t = \tau_k$ for all $k \in \{0, 1, ..., N_T^i\}$ and compare (66) to (70). Furthermore, observe that $\widehat{\Theta}_t^i - \widehat{\Theta}_{t-}^i = -\frac{\lambda_{1,t}}{\lambda_{2,t}}(\widehat{\Pi}_t^i - \widehat{\Pi}_{t-}^i)$. It then follows that, $\forall k \in \{0, 1, ..., N_T^i\}$,

$$\widehat{X}_t^i = \widehat{X}_{t-}^i + \begin{bmatrix} 1 \\ -\frac{\lambda_{1,t}}{\lambda_{2,t}} \end{bmatrix} o_t^i \frac{\Delta n_t^i}{\sigma_S^2} \widehat{Y}_t^i \equiv \widehat{X}_{t-}^i + \omega_t o_t^i \frac{\Delta n_t^i}{\sigma_S^2} \widehat{Y}_t^i, \quad \forall t = \tau_k \tag{74}$$

and

$$\frac{1}{o_t^i} = \frac{1}{o_{t-}^i} + \frac{\Delta n_t^i}{\sigma_S^2}, \quad \forall t = \tau_k.$$

The proof of the more general expressions in (71) and (72) uses the law of conditional expectations for Gaussian variables (see, e.g., Jazwinski (1970), Theorem 7.1). The distribution of the filter innovation $\widehat{Y}_t^i$ in (73) follows from the result of Lemma A.1 that the aggregate signal Y_t^i is Gaussian conditional on the filtration $\mathscr{F}_{t-}^i$ and Δn_t^i and from the observation that

$$\mathbb{E}\left[Y_t^i - \mathbb{E}\left[Y_t^i | \mathscr{F}_{t-}^i, \Delta n_t^i\right] \middle| \mathscr{F}_{t-}^i, \Delta n_t^i\right] = 0$$

and

$$\mathbb{V}\left[Y_t^i - \mathbb{E}\left[Y_t^i | \mathscr{F}_{t-}^i, \Delta n_t^i\right] \middle| \mathscr{F}_{t-}^i, \Delta n_t^i\right] = \mathbb{V}\left[\Pi + \frac{\sigma_S}{\Delta n_t^i} \sum_{j=1}^{\Delta n_t^i} \epsilon_{j+n_{t-}^i}^i \middle| \mathscr{F}_{t-}^i, \Delta n_t^i\right] = o_{t-}^i + \frac{\sigma_S^2}{\Delta n_t^i}.$$

□

Theorem A.2 implies that, at time $t = \tau_k$ for any $k \in \{0, 1, ..., N_T^i\}$, the conditional mean $\widehat{X}_t^i$ and the conditional variance o_t^i change discontinuously by $\Delta \widehat{X}_t^i \equiv \widehat{X}_t^i - \widehat{X}_{t-}^i = \omega_t o_t^i \frac{\Delta n_t^i}{\sigma_S^2} \widehat{Y}_t^i$ and $\Delta o_t^i \equiv o_t^i - o_{t-}^i = -\frac{\Delta n_t^i}{\sigma_S^2} o_t^i o_{t-}^i$, respectively. This result and the result of Theorem A.1 imply in turn that, at any time $t \in (0, T)$, $\widehat{X}_t^i$ and o_t^i have the dynamics highlighted in Lemma A.2.

Lemma A.2. *At any time $t \in (0, T)$, the conditional mean $\widehat{X}_t^i$ and the conditional variance o_t^i have dynamics*

$$d\widehat{X}_t^i = a\widehat{X}_{t-}^i\, dt + \begin{bmatrix} o_{t-}^i k_t \\ \left(\sigma_\Theta - o_{t-}^i \frac{\lambda_{1,t}}{\lambda_{2,t}} k_t\right) \end{bmatrix} d\widehat{B}_t^i + \omega_t o_t^i \frac{\Delta n_t^i}{\sigma_S^2} \widehat{Y}_t^i\, dN_t^i, \quad \widehat{X}_0^i = O_{0-} A_0^\top \left(A_0 O_{0-} A_0^\top + B_0 B_0^\top\right)^{-1} \Xi_0^i \tag{75}$$

$$do_t^i = -k_t^2 (o_{t-}^i)^2\, dt - \frac{\Delta n_t^i}{\sigma_S^2} o_t^i o_{t-}^i\, dN_t^i, \quad o_0^i = \left(\frac{1}{\sigma_\Pi^2} + \frac{n_0^i}{\sigma_S^2} + \left(\frac{\lambda_{1,0}}{\lambda_{2,0}}\right)^2 \frac{1}{\sigma_\Theta^2}\right)^{-1} \tag{76}$$

where $\omega_t \equiv \begin{bmatrix} 1 & -\frac{\lambda_{1,t}}{\lambda_{2,t}} \end{bmatrix}^\top$ *and*

$$O_{0-} = \begin{bmatrix} \sigma_\Pi^2 & 0 \\ 0 & \sigma_\Theta^2 \end{bmatrix}, \quad A_0 = \begin{bmatrix} \lambda_{1,0} & \lambda_{2,0} \\ 1 & 0 \end{bmatrix}, \quad B_0 = \begin{bmatrix} 0 \\ \frac{\sigma_S}{\sqrt{n_0^i}} \end{bmatrix}, \quad \Xi_0^i = \begin{bmatrix} \xi_0 \\ Y_0^i \end{bmatrix}. \tag{77}$$

Proof. To obtain the initial conditions in (75), notice that managers start with priors $\widehat{\Pi}_{0-}^i \sim \mathscr{N}(0, \sigma_\Pi^2)$ and $\widehat{\Theta}_{0-}^i \sim \mathscr{N}(0, \sigma_\Theta^2)$ and immediately observe the vector Ξ_0^i in (77). They thus initially update their expectations by applying Theorem A.2 using the matrices in (77). The remainder follows directly from Theorem A.1 and A.2 and the simplifications above. □

I now define the *common* information set $\mathscr{F}_t^c$ at time t as

$$\mathscr{F}_t^c = \sigma\left((\xi_s) : 0 \le s \le t\right), \quad 0 \le t < T.$$

Notice that the filtration $\mathscr{F}_t^c$ is Brownian and therefore left-continuous (Karatzas and Shreve (1988), Problem 7.1), i.e., there is no "surprise" in the common information flow. Accordingly, in contrast to the dynamics in (75), the dynamics of the conditional mean $\widehat{X}_t^c = \mathbb{E}\left[X_t | \mathscr{F}_t^c\right]$ and the conditional variance $o_t^c = \mathbb{E}\left[(\Pi - \widehat{\Pi}_t^c)^2 \middle| \mathscr{F}_t^c\right]$ are continuous, as shown in Corollary 5.

Corollary 5. *The conditional mean $\widehat{X}_t^c$ with respect to the filtration $\mathscr{F}_t^c$ has dynamics*

$$d\widehat{X}_t^c = a\widehat{X}_t^c\, dt + \begin{bmatrix} o_t^c k_t \\ \left(\sigma_\Theta - o_t^c \frac{\lambda_{1,t}}{\lambda_{2,t}} k_t\right) \end{bmatrix} d\widehat{B}_t^c, \quad \widehat{X}_0^c = O_{0-} A_{1,0}^\top \left(A_{1,0} O_{0-} A_{1,0}^\top\right)^{-1} \xi_0 \tag{78}$$

and the conditional variance o_t^c is given by the solution to the Ricatti equation

$$do_t^c = -k_t^2 (o_t^c)^2 \, dt, \quad o_0^c = \left(\frac{1}{\sigma_\Pi^2} + \left(\frac{\lambda_{1,0}}{\lambda_{2,0}} \right)^2 \frac{1}{\sigma_\Theta^2} \right)^{-1} \tag{79}$$

where $A_{1,0} = \begin{bmatrix} \lambda_{1,0} & \lambda_{2,0} \end{bmatrix}$ and where the filter innovation $\widehat{B}_t^c$ satisfying

$$d\widehat{B}_t^c = (B_{1,t} B_{1,t}^\top)^{-\frac{1}{2}} (d\xi_t - A_{1,t} \widehat{X}_t^c \, dt) \tag{80}$$

is a $\widehat{\mathbb{P}}^c$−Brownian motion with respect to the filtration $\mathscr{F}_t^c$.

Proof. Given that, for all $t \in (0, T)$, the filtration $\mathscr{F}_t^c$ is continuous, the proof follows directly from Theorem A.1 and from simplifications based on (58). The initial conditions follow as a special case of those in Lemma A.2. □

Remark 2. *Relating the common conditional variance o^c in (82) to agent i's conditional variance o^i in (75) shows that $o_t^i \equiv o_t(n^i)$ is a function of her number of signals n^i and time t only:*

$$o_t(n^i) = \mathbb{E}\left[(\Pi - \widehat{\Pi}_t^i)^2 \middle| \mathscr{F}_t^i ; n_t^i = n^i \right] = \left(\frac{1}{o_t^c} + \frac{1}{\sigma_S^2} \sum_{s \le t} \Delta n_s^i \Delta N_s^i \right)^{-1} = \left(\frac{1}{o_t^c} + \frac{n^i}{\sigma_S^2} \right)^{-1} \tag{81}$$

where, applying Ito's lemma to $(o_t^c)^{-1}$ using (79), the common precision is given explicitly by

$$\frac{1}{o_t^c} = \frac{1}{\sigma_\Pi^2} + \left(\frac{\lambda_{1,0}}{\lambda_{2,0}} \right)^2 \frac{1}{\sigma_\Theta^2} + \int_0^t k_s^2 \, ds. \tag{82}$$

I conclude by characterizing the dynamics of stock returns with respect to $\mathscr{F}^i$. Let

$$\mathrm{d}Q_t = \mathrm{d}P_t - rP_t \mathrm{d}t + \mathbf{1}_{\{t=T\}} \Delta P_T \tag{83}$$

denote the instantaneous excess return on one share of the stock and let

$$\Delta P_T = \Pi - P_{T-} = (1 - \lambda_{1,T-})(\Pi - \widehat{\Pi}_{T-}^c) - \lambda_{2,T-} \Theta_T \tag{84}$$

represent a price discontinuity at time T when the price reaches $P_T = \Pi$. Furthermore, let $\Delta_t \equiv \Pi - \widehat{\Pi}_t^c$ and define the vector $\Psi_t = (\Delta_t, \Theta_t)^\top$ and its conditional expectation

$$\Psi_t^i \equiv \mathbb{E}[\Psi_t | \mathscr{F}_t^i] = (\Delta_t^i, \widehat{\Theta}_t^i)^\top. \tag{85}$$

Then, the instantaneous excess return with respect to $\mathscr{F}^i$ is fully characterized by (Ψ^i, n^i), as Lemma A.3 shows.

Lemma A.3. *The instantaneous excess return dQ with respect to agent i's filtration $\mathscr{F}_t^i$ satisfies*

$$dQ_t = \begin{bmatrix} \lambda_{1,t}' + (1 - \lambda_{1,t}) o_t^c k_t^2 & \lambda_{2,t}' - a_\Theta \lambda_{2,t} \end{bmatrix} \Psi_t^i dt + (\lambda_{2,t} \sigma_\Theta + (1 - \lambda_{1,t}) o_t^c k_t) d\widehat{B}_t^i + \mathbf{1}_{\{t=T\}} \Delta Q_T \tag{86}$$

$$\equiv A_{Q,t} \Psi_t^i dt + B_{Q,t} d\widehat{B}_t^i + \mathbf{1}_{\{t=T\}} \Delta Q_T \tag{87}$$

where the conditional expectation Ψ_t^i, as defined in (85)*, is a Gaussian process with dynamics*

$$d\Psi_t^i = \begin{bmatrix} -k_t^2 o_t^c & 0 \\ 0 & -a_\Theta \end{bmatrix} \Psi_{t-}^i dt + \begin{bmatrix} (o_t(n_{t-}^i) - o_t^c)k_t \\ \sigma_\Theta - \frac{\lambda_{1,t}}{\lambda_{2,t}} o_t(n_{t-}^i)k_t \end{bmatrix} d\widehat{B}_t^i + \omega_t o_t(n_t^i) \frac{\Delta n_t^i}{\sigma_S^2} \widehat{Y}_t^i dN_t^i \tag{88}$$

$$\equiv A_{\Psi,t}\Psi_{t-}^i dt + B_{\Psi,t}(n_{t-}^i) d\widehat{B}_t^i + \sigma_t(n_{t-}^i, \Delta n_t^i)\widehat{Y}_t^i dN_t^i. \tag{89}$$

Furthermore, letting $\lambda_t \equiv \begin{bmatrix} 1-\lambda_{1,t} & -\lambda_{2,t} \end{bmatrix}^\top$, the excess return ΔQ_T at the liquidation date satisfies

$$\Delta Q_T \equiv \lambda_{T-}^\top \Psi_{T-}^i. \tag{90}$$

Equations (86) *and* (90) *in turn imply that the Markov process $(\Psi_t^i, n_t^i)_{t\geq 0}$ fully characterizes instantaneous excess returns under $\mathscr{F}_t^i$.*

Proof. To derive the dynamics in (86) and (88), I need to relate the probability measures $\widehat{\mathbb{P}}^c$ and $\widehat{\mathbb{P}}^i$ under the filtration $\mathscr{F}^i$: define the Radon-Nikodym derivative Z_t of $\widehat{\mathbb{P}}^c$ with respect to $\widehat{\mathbb{P}}^i$ under $\mathscr{F}^i$ as

$$Z_t = \left.\frac{\mathrm{d}\widehat{\mathbb{P}}^i}{\mathrm{d}\widehat{\mathbb{P}}^c}\right|_{\mathscr{F}_t^i} = \exp\left(-\frac{1}{2}\int_0^t (k_s\Delta_s^i)^2 \mathrm{d}s + \int_0^t k_s\Delta_s^i \mathrm{d}\widehat{B}_s^c\right) \tag{91}$$

where $\Delta^i \equiv \widehat{\Pi}^i - \widehat{\Pi}^c$. The process Z defines a change of measure between $\widehat{\mathbb{P}}^c$ and $\widehat{\mathbb{P}}^i$ under $\mathscr{F}^i$, a result I establish in Theorem A.3.

Theorem A.3. *Let $(\widehat{B}^c)_{t\geq 0}$ be a $\widehat{\mathbb{P}}^c-$Brownian motion with differentials as in* (80)*, and let $(\mathscr{F}^i)_{t\geq 0}$, as defined in* (80)*, be the filtration for this Brownian motion. Let $(k_t\Delta_t^i)_{t\geq 0}$ be an adapted process. Then, $\mathbb{E}^{\widehat{\mathbb{P}}^c}[Z] = 1$ and the process $(\widehat{B}^i)_{t\geq 0}$ satisfying*

$$\widehat{B}_t^i = \widehat{B}_t^c - \int_0^t k_s\Delta_s^i ds \tag{92}$$

is a $\widehat{\mathbb{P}}^i-$Brownian motion with respect to $\mathscr{F}_t^i$.

Proof. First, observe that the Brownian motion $\widehat{B}^c$ is adapted to $\mathscr{F}^c$ and, therefore, to $\mathscr{F}^i \supset \mathscr{F}^c$. Second, combine (57) and (80) and obtain

$$\mathrm{d}\widehat{B}_t^i = \mathrm{d}\widehat{B}_t^c - \frac{1}{\lambda_{2,t}\sigma_\Theta}(\lambda_{1,t}'\Delta_t^i + (\lambda_{2,t}' - a_\Theta\lambda_{2,t})(\widehat{\Theta}_t^i - \widehat{\Theta}_t^c))\mathrm{d}t. \tag{93}$$

Since $\xi_t \in \mathscr{F}_t^c \subset \mathscr{F}_t^i$, it follows that $\widehat{\Theta}_t^i - \widehat{\Theta}_t^c = -\frac{\lambda_{1,t}}{\lambda_{2,t}}\Delta_t^i$, which, substituted in (93) and using (61) gives (92). Third, for $E^{\widehat{\mathbb{P}}^c}[Z] = 1$ to hold, i.e., for $\widehat{\mathbb{P}}^i$ to be absolutely continuous with respect to $\widehat{\mathbb{P}}^c$ under $\mathscr{F}^i$, the Radon-Nikodym derivative in (91) must be a martingale. A sufficient condition under which Z is a martingale is the Novikov condition (Karatzas and Shreve (1988), Proposition 5.12):

$$\mathbb{E}^{\widehat{\mathbb{P}}^c}\left[\exp\left(\frac{1}{2}\int_0^T (k_t\Delta_t^i)^2 \mathrm{d}t\right)\right] < \infty, \quad 0 \leq T \leq \infty \tag{94}$$

where the process Δ^i under $\widehat{\mathbb{P}}^c$ satisfies

$$\mathrm{d}\Delta_t^i = -o_t(n^i)k_t^2\Delta_t^i\mathrm{d}t + k_t(o_t(n^i) - o_t^c)\mathrm{d}\widehat{B}_t^c. \tag{95}$$

Observing that the process in (95) is Gaussian, Example 3 (a) (Liptser and Shiryaev (2000), p. 233) shows that the Novikov condition in (94) boils down to

$$\sup_{t\le T}|k_t|\mathbb{E}^{\widehat{\mathbb{P}}^c}\left[\left|\Delta_t^i\right|\right] < \infty, \quad \sup_{t\le T}k_t^2\mathbb{V}^{\widehat{\mathbb{P}}^c}\left[\Delta_t^i\right] < \infty. \tag{96}$$

In this setup, the conditions in (96) simplify to

$$\sup_{t\le T}|k_t|\int_0^t k_s^2\mathrm{d}s < \infty, \quad \sup_{t\le T}k_t^2\int_0^t k_s^2\mathrm{d}s < \infty. \tag{97}$$

Using (61) and anticipating on the equilibrium result of Lemma C.2, the two conditions in (97) are equivalent to requiring that the function $\phi(\cdot)$ be continuous, which it is by assumption. Theorem A.3 then follows from Girsanov theorem (Karatzas and Shreve (1988), Theorem 5.1). □

Using the change of measure of Theorem A.3 and the dynamics in (78), the dynamics of common expectations under $\widehat{\mathbb{P}}^i$ with respect to the filtration $\mathscr{F}^i$ satisfy

$$\mathrm{d}\widehat{X}_t^c = \left(a\widehat{X}_t^c + \begin{bmatrix} o_t^c k_t^2 \\ k_t\left(\sigma_\Theta - o_t^c\frac{\lambda_{1,t}}{\lambda_{2,t}}k_t\right)\end{bmatrix}\Delta_t^i\right)\mathrm{d}t + \begin{bmatrix} o_t^c k_t \\ \left(\sigma_\Theta - o_t^c\frac{\lambda_{1,t}}{\lambda_{2,t}}k_t\right)\end{bmatrix}\mathrm{d}\widehat{B}_t^i. \tag{98}$$

Applying Ito's lemma and using (75) and (98), it then follows that Δ^i has dynamics

$$\mathrm{d}\Delta_t^i = -o_t^c k_t^2\Delta_{t-}^i\mathrm{d}t + \left(o_t(n_{t-}^i) - o_t^c\right)k_t\mathrm{d}\widehat{B}_t^i + o_t(n_t^i)\frac{\Delta n_t^i}{\sigma_S^2}\widehat{Y}_t^i\mathrm{d}N_t^i. \tag{99}$$

Reorganizing the dynamics in (75) and (99) and using (81), I obtain the dynamics in (88). Inspection of (88) then reveals that $(\Psi_t^i, n_t^i)_{t\ge 0}$ is a Markov process. Furthermore, set $r = 0$ and observe that $Q_t \equiv \lambda_{1,t}\widehat{\Pi}_t^i + \lambda_{2,t}\widehat{\Theta}_t^i + (1-\lambda_{1,t})\widehat{\Pi}_t^c$ under $\widehat{\mathbb{P}}^i$. Applying Ito's lemma to this expression and substituting the dynamics of $\widehat{\Pi}_t^c$ with respect to $\mathscr{F}_t^i$ (in (98)) shows that excess returns $\mathrm{d}Q$ satisfy

$$\mathrm{d}Q_t = ((\lambda'_{1,t} + (1-\lambda_{1,t})o_t^c k_t^2)(\widehat{\Pi}_t^i - \widehat{\Pi}_t^c) + (\lambda'_{2,t} - a_\Theta\lambda_{2,t})\widehat{\Theta}_t^i)\mathrm{d}t + (\lambda_{2,t}\sigma_\Theta + (1-\lambda_{1,t})o_t^c k_t)\mathrm{d}\widehat{B}_t^i \tag{100}$$

under agent i's filtration $\mathscr{F}_t^i$. Rewriting the drift of $\mathrm{d}Q$ in terms of the Gaussian process Ψ^i and taking into account (83), equation (86) follows. Finally, taking expectations of the price discontinuity in (84) with respect to $\mathscr{F}_{T-}^i$, I get

$$\mathbb{E}[\Delta Q_T|\mathscr{F}_{T-}^i] = \mathbb{E}[\Delta P_T|\mathscr{F}_{T-}^i] = \begin{bmatrix} 1-\lambda_{1,T-} & -\lambda_{2,T-}\end{bmatrix}\mathbb{E}[\Psi_T|\mathscr{F}_{T-}^i]$$

and equation (90) follows. □

Appendix B. Proof of Proposition 2 (optimal demands)

In this appendix, I solve the portfolio optimization problem of an agent who collects ideas according to the mechanism of Section 2.1 and whose expectations satisfy the dynamics of Proposition 1. I show that her optimal asset demand has an explicit solution.

Considering the optimization problem defined in (5) and using Lemma A.3, an investor's value function depends exclusively on (W, Ψ, n, t). Accordingly, let the value function of agent i at time t be

$$J(W^i, \Psi^i, n^i, t) = \max_{\theta^i} \mathbb{E}\left[-\exp\left(-\gamma W_T^i\right)\middle|\, \mathscr{F}_t^i; W_t^i = W^i, \Psi_t^i = \Psi^i, n_t^i = n^i\right] \tag{101}$$

$$\text{s.t. } \mathrm{d}W_t^i = rW_t^i\mathrm{d}t + \theta_t^i \mathrm{d}Q_t. \tag{102}$$

The value function in (101) must satisfy the Hamilton-Jacobi-Bellman (HJB) equation

$$0 = \max_{\theta^i}\left\{J_W A_Q \Psi^i \theta^i + \frac{1}{2} J_{WW} B_Q^2 (\theta^i)^2 + B_Q B_\Psi(n^i)^\top J_{W\Psi}\theta^i\right\} + J_t + J_\Psi^\top A_\Psi \Psi^i + \frac{1}{2}\mathrm{tr}(J_{\Psi\Psi} B_\Psi(n^i) B_\Psi(n^i)^\top) \tag{103}$$

$$+ \eta(n^i)\mathbb{E}^{\mathscr{L}_t(\widehat{Y}^i, \Delta n^i)}\left[J(W^i, \Psi^i + \sigma(n^i, \Delta n^i)\widehat{Y}^i, n^i + \Delta n^i, t) - J(W^i, \Psi^i, n^i, t)\right], \tag{104}$$

where the coefficients A and B are defined in Lemma A.3, and the boundary condition

$$J(W^i, \Psi^i, n^i, T-) = \max_{\theta^i} \mathbb{E}\left[-\exp\left(-\gamma(W_{T-}^i + \theta_{T-}^i \Delta Q_T)\right)\middle|\, \mathscr{F}_{T-}^i; W_{T-}^i = W^i, \Psi_{T-}^i = \Psi^i, n_{T-}^i = n^i\right]. \tag{105}$$

The last term in (103) represents the probability, $\mathbb{P}[\mathrm{d}N_t^i = 1 | n_{t-}^i = n] = \eta(n)\mathrm{d}t$, that agent i collects new ideas in the time interval $(t, t + \mathrm{d}t)$ times the expected change in her value function, which involves the joint distribution $\mathscr{L}_t$ of her filter innovation, $\widehat{Y}^i$, and her number, Δn^i, of new signals given that new ideas are collected at time t. The boundary condition in (105) is a static optimization problem, which accounts for the price discontinuity at time T (see Lemma A.3).

Solving the maximization problem in (103) yields the optimal portfolio policy

$$\theta_t^i \equiv \theta_t(\Psi^i, n^i) = -\frac{J_W A_Q \Psi^i + B_Q B_\Psi(n^i)^\top J_{W\Psi}}{J_{WW} B_Q^2}. \tag{106}$$

The second-order condition for optimality is $J_{WW} < 0$. Substituting the optimal portfolio policy into (103), I obtain a partial integro-differential equation (PIDE) for J:

$$0 = J_t + J_\Psi^\top A_\Psi \Psi^i + \frac{1}{2}\mathrm{tr}(J_{\Psi\Psi} B_\Psi(n^i) B_\Psi(n^i)^\top) - \frac{1}{2}\frac{(J_W A_Q \Psi^i + B_Q B_\Psi(n^i)^\top J_{W\Psi})^2}{J_{WW} B_Q^2} \tag{107}$$

$$+ \eta(n^i)\mathbb{E}^{\mathscr{L}_t(\widehat{Y}^i, \Delta n^i)}\left[J(W^i, \Psi^i + \sigma(n^i, \Delta n^i)\widehat{Y}^i, n^i + \Delta n^i, t) - J(W^i, \Psi^i, n^i, t)\right]. \tag{108}$$

The PIDE in (107) decouples into an integro-differential equation (IDE), a system of differential equations in the time dimension and a system of functional equations in the number-of-signals dimension, as shown in Theorem B.1.

Theorem B.1. *The PIDE in* (107) *with boundary condition in* (105) *has a solution of the form*

$$J(W, \Psi, n, t) = -\exp\left(-\gamma W - u_t(n) - \frac{1}{2}\left(\Psi^\top R_t(n) + R_t(n)^\top \Psi + \Psi^\top M_t(n)\Psi\right)\right) \tag{109}$$

where, at any time $t \in [0,T)$ and for any $n \in \mathbb{N}$, the scalar coefficient $u_t(n)$ satisfies the IDE

$$\frac{d}{dt}u_t(n) = \eta_t(n)\sum_{m\in\mathbb{N}} \pi_t(m)\,|C_t(n,m)|^{\frac{1}{2}} \exp\left(\begin{array}{c} -(u_t(n+m)-u_t(n)) \\ +\frac{1}{2}R_t(n+m)^\top C_t(n,m)\Sigma_t(n,m)R_t(n+m) \end{array}\right) \tag{110}$$

$$-\frac{1}{2}tr(M_t(n)B_{\Psi,t}(n)B_{\Psi,t}(n)^\top) - \eta_t(n), \quad u_{T-}(n) = 0 \tag{111}$$

the $(2\times 1)-$vector coefficient $R_t(n)$ satisfies the differential equation

$$\dot{R}_t(n) = \left(\frac{B_{\Psi,t}(n)A_{Q,t}}{B_{Q,t}} - A_{\Psi,t}\right)^\top R_t(n), \quad R_{T-}(n) = \mathbf{0} \tag{112}$$

along with the functional equation

$$R_t(n) = C_t(n,m)R_t(n+m), \quad \forall m \in \mathbb{N} \tag{113}$$

and the $(2\times 2)-$symmetric matrix coefficient $M_t(n)$ satisfies the differential equation

$$\dot{M}_t(n) = -\frac{A_{Q,t}^\top A_{Q,t}}{B_{Q,t}^2} + M_t(n)\left(\frac{B_{\Psi,t}(n)A_{Q,t}}{B_{Q,t}} - A_{\Psi,t}\right) + \left(\frac{B_{\Psi,t}(n)A_{Q,t}}{B_{Q,t}} - A_{\Psi,t}\right)^\top M_t(n) \tag{114}$$

with boundary condition

$$M_{T-}(n) = (o_{T-}(n))^{-1}\left[\begin{array}{cc} (1-\lambda_{1,T_-})^2 & -\lambda_{2,T_-}(1-\lambda_{1,T_-}) \\ -\lambda_{2,T_-}(1-\lambda_{1,T_-}) & \lambda_{2,T_-}^2 \end{array}\right] \equiv (o_{T-}(n))^{-1}\Lambda_{T-} \tag{115}$$

along with the functional equation

$$M_t(n) = C_t(n,m)M_t(n+m), \quad \forall m \in \mathbb{N}. \tag{116}$$

The $(2\times 2)-$symmetric matrix C is defined as

$$C_t(n,m) = (\mathbf{I} + M_t(n+m)\Sigma_t(n,m))^{-1}, \tag{117}$$

the $(2\times 2)-$symmetric matrix Σ is defined as

$$\Sigma_t(n,m) = \mathbb{E}\left[\Delta\widehat{X}_t^i(\Delta\widehat{X}_t^i)^\top\middle|\mathscr{F}_{t-}^i; n_{t-}^i = n, \Delta n_t^i = m\right] = \sigma_t(n,m)\sigma_t(n,m)^\top\mathbb{V}\left[Y_t^i\middle|\mathscr{F}_t^i; n_{t-}^i = n, \Delta n_t^i = m\right] \tag{118}$$

$$= \Omega_t\frac{m}{\sigma_S^2}o_t(n)o_t(n+m) \tag{119}$$

and the $(2\times 2)-$symmetric matrix Ω is defined in (58).

Proof. The proof is organized in two parts. I first show that the functional equations in (113) and (116) are necessary and sufficient conditions for the ansatz in (109) and the differential equations in (110), (112) and (114) to hold. Second, I show that the differential equations in (110), (112) and (114) satisfy the functional equations in (113) and (116), which eventually validates the ansatz in (109).

First, substituting the ansatz in (109) into the boundary condition in (105) and using Lemma

A.3, I obtain

$$J(W,\Psi,n,T-) = \max_{\theta^i} \mathbb{E}\left[-\exp\left(-\gamma(W^i_{T-} + \theta^i_{T-}\lambda^\top_{T-}\Psi_T)\right)\Big| \mathscr{F}^i_{T-}; W^i_{T-} = W, \Psi^i_{T-} = \Psi, n^i_{T-} = n\right] \tag{120}$$

$$= \max_{\theta^i} -\exp\left(-\gamma\left(W + \theta^i_{T-}\lambda^\top_{T-}\Psi - \frac{1}{2}\gamma(\theta^i_{T-})^2\lambda^\top_{T-}\Omega_{T-}\lambda_{T-}o^i_{T-}\right)\right) \tag{121}$$

where the second equality uses the Laplace transform for Gaussian variables and the result of Lemma A.3 that Ψ is Gaussian with respect to $\mathscr{F}^i$. Solving the optimization problem in (120) and using that $\lambda^\top_{T-}\Omega_{T-}\lambda_{T-} = 1$ and (81) yields the optimal portfolio policy

$$\theta^i_{T-} \equiv \theta^i(\Psi,n,T-) = \frac{\mathbb{E}\left[\Delta P_T|\mathscr{F}^i_{T-}; \Psi^i_{T-} = \Psi, n^i_{T-} = n\right]}{\gamma\mathbb{V}\left[\Delta P_T|\mathscr{F}^i_{T-}; \Psi^i_{T-} = \Psi, n^i_{T-} = n\right]} = \frac{1}{\gamma}\left(o_{T-}(n)\right)^{-1}\lambda^\top_{T-}\Psi. \tag{122}$$

Substituting the optimal portfolio policy into (120) and using that $\lambda_{T-}\lambda^\top_{T-} = \Lambda_{T-}$, I obtain

$$J(W,\Psi,n,T-) = -\exp\left(-\gamma W - \frac{1}{2}\left(o_{T-}(n)\right)^{-1}\Psi^\top\Lambda_{T-}\Psi\right). \tag{123}$$

Comparing this expression with the ansatz in (109) evaluated at time $T-$, I obtain the boundary conditions in (110), (112) and (114). Second, substituting the ansatz in (109) into the PIDE in (107) and simplifying, I obtain

$$0 = -\frac{\mathrm{d}}{\mathrm{d}t}u_t(n) - \frac{1}{2}\mathrm{tr}(M_t(n)B_{\Psi,t}(n)B_{\Psi,t}(n)^\top) + \eta_t(n)\left(F_t(\Psi,n) - 1\right) \tag{124}$$

$$+ \frac{1}{2}\Psi^\top\left(\left(\frac{B_{\Psi,t}(n)A_{Q,t}}{B_{Q,t}} - A_{\Psi,t}\right)^\top R_t(n) - \dot{R}_t(n)\right) + \frac{1}{2}\left(R_t(n)^\top\left(\frac{B_{\Psi,t}(n)A_{Q,t}}{B_{Q,t}} - A_{\Psi,t}\right) - \dot{R}_t(n)^\top\right)\Psi \tag{125}$$

$$+ \frac{1}{2}\Psi^\top\left(-\frac{A^\top_{Q,t}A_{Q,t}}{B^2_{Q,t}} + M_t(n)\left(\frac{B_{\Psi,t}(n)A_{Q,t}}{B_{Q,t}} - A_{\Psi,t}\right) + \left(\frac{B_{\Psi,t}(n)A_{Q,t}}{B_{Q,t}} - A_{\Psi,t}\right)^\top M_t(n) - \dot{M}_t(n)\right)\Psi \tag{126}$$

where $F_t(\Psi,n)$ denotes the expectation

$$F_t(\Psi,n) \equiv \mathbb{E}^{\mathscr{L}_t}\left[\exp\left(-H_t(\Psi,n,\Delta n) - \sigma_t(n,\Delta n)^\top G_t(\Psi,n+\Delta n)\widehat{Y} - \frac{\sigma_t(n,\Delta n)^\top M_t(n+\Delta n)\sigma_t(n,\Delta n)\widehat{Y}^2}{2}\right)\right]. \tag{127}$$

Letting $\Delta u_t(n,\Delta n) \equiv u_t(n+\Delta n) - u_t(n)$, $\Delta R_t(n,\Delta n) \equiv R_t(n+\Delta n) - R_t(n)$ and $\Delta M_t(n,\Delta n) \equiv M_t(n+\Delta n) - M_t(n)$, the scalar H is defined as

$$H_t(\Psi,n,\Delta n) = \Delta u_t(n,\Delta n) + \frac{1}{2}\left(\Psi^\top\Delta R_t(n,\Delta n) + \Delta R_t(n,\Delta n)^\top\Psi\right) + \frac{1}{2}\Psi^\top\Delta M_t(n,\Delta n)\Psi \tag{128}$$

and the $(2\times 1)-$vector G is defined as

$$G_t(\Psi,n+\Delta n) = R_t(n+\Delta n) + M_t(n+\Delta n)\Psi. \tag{129}$$

To compute the expectation in (127), I need to determine the joint distribution $\mathscr{L}_t$ of $\widehat{Y}$ and Δn conditional on collecting new ideas at time t. Applying Bayes' rule and using the result of Lemma A.2, conditional on collecting new ideas in $(t, t + \mathrm{d}t)$, the probability density function l_t of $(\widehat{Y}, \Delta n_t) \sim \mathscr{L}_t$ satisfies

$$l_t(y, m; n) \equiv \mathbb{P}[\widehat{Y} \in \mathrm{d}y | \mathscr{F}^i_{t-}; \Delta n^i_t = m, n^i_{t-} = n] \times \mathbb{P}[\Delta n^i_t = m | \mathrm{d}N^i_t = 1, n^i_{t-} = n] \tag{130}$$

$$= \pi_t(m; n) \left(2\pi \left(o_t(n) + \frac{\sigma_S^2}{m}\right)\right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\left(o_t(n) + \frac{\sigma_S^2}{m}\right)^{-1} y^2\right) \mathrm{d}y. \tag{131}$$

The expectation in (127) can then be computed explicitly using the result of Lemma B.1.

Lemma B.1. *Let H and v be scalars, G and σ be $(2 \times 1)-$vectors, M be a $(2 \times 2)-$positive semi-definite matrix and let $y \sim \mathscr{N}(0, v)$. Furthermore, define $\Sigma = \sigma^\top \sigma v$ and assume that $\mathbf{I} + \Sigma M$ is positive definite. Then,*

$$\mathbb{E}\left[\exp\left(-H - \sigma^\top G y - \frac{1}{2}\sigma^\top M \sigma y^2\right)\right] = |\mathbf{I} + \Sigma M|^{-\frac{1}{2}} \exp\left(\frac{1}{2} G^\top (\mathbf{I} + \Sigma M)^{-1} \Sigma G - H\right). \tag{132}$$

Proof. Lemma B.1 follows from the moment-generating function of a noncentral chi-square distribution. □

Using Lemma B.1, it follows that

$$F_t(\Psi, n) = \sum_{m \in \mathbb{N}} \pi_t(m; n) \frac{\exp\left(\begin{array}{c} \frac{1}{2} G_t(\Psi, n+m)^\top (\mathbf{I} + \Sigma_t(n,m) M_t(n+m))^{-1} \Sigma_t(n,m) G_t(\Psi, n+m) \\ -H_t(\Psi, n, m) \end{array}\right)}{|\mathbf{I} + \Sigma_t(n,m) M_t(n+m)|^{\frac{1}{2}}} \tag{133}$$

$$= \sum_{m \in \mathbb{N}} \pi_t(m; n) \frac{\exp\left(\begin{array}{c} -(u_t(n+m) - u_t(n)) + \frac{1}{2} R_t(n+m)^\top C_t(n,m) \Sigma_t(n,m) R_t(n+m) \\ -\frac{1}{2}\Psi^\top (C_t(n,m) R_t(n+m) - R_t(n)) \\ -\frac{1}{2}(C_t(n,m) R_t(n+m) - R_t(n))^\top \Psi \\ -\frac{1}{2}\Psi^\top (C_t(n,m) M_t(n+m) - M_t(n)) \Psi \end{array}\right)}{|\mathbf{I} + \Sigma_t(n,m) M_t(n+m)|^{\frac{1}{2}}} \tag{134}$$

where the last equality follows from substituting the coefficient H in (128) and the coefficient G in (129), simplifying, and using Woodbury matrix identity

$$C_t(n, m) \equiv \mathbf{I} - M_t(n+m)(\mathbf{I} + \Sigma_t(n,m) M_t(n+m))^{-1} \Sigma_t(n,m) = (\mathbf{I} + M_t(n+m) \Sigma_t(n,m))^{-1}. \tag{135}$$

Finally, for the solution to the PIDE in (107) to have the form of the ansatz in (109), it must be that $F_t(\Psi, n) \equiv F_t(n)$ at any time $t \in [0, T)$. Clearly, $F_t(\Psi, n) \equiv F_t(n)$ if and only if the functional equations in (113) and (116) hold true. Indeed, substituting (113) and (116) into (133) and simplifying, I obtain

$$F_t(\Psi, n) = \sum_{m \in \mathbb{N}} \pi_t(m; n) \, |C_t(n,m)|^{\frac{1}{2}} \exp\left(\begin{array}{c} -(u_t(n+m) - u_t(n)) \\ +\frac{1}{2} R_t(n+m)^\top C_t(n,m) \Sigma_t(n,m) R_t(n+m) \end{array}\right) \tag{136}$$

$$\equiv F_t(n) \tag{137}$$

where I use that $|\mathbf{I} + \Sigma_t(n,m)M_t(n+m)|^{-1} = |C_t(n+m)|$. Plugging this expression back into (124) and separating variables yields the equations in (110), (112) and (114), which concludes the first part of the proof.

Given that the functional equations in (113) and (116) are necessary and sufficient conditions for the ansatz in (109) to hold, I still need to show that the differential equations in (112) and (114) satisfy these functional equations. First, substituting (116) into the boundary condition in (115), I obtain

$$C_{T-}(n,m)M_{T-}(n+m) = (o_{T-}(n))^{-1}\,\Lambda_{T-}. \tag{138}$$

Multiplying both sides of this expression by $C_{T-}(n,m)^{-1}$ and simplifying shows that

$$M_{T-}(n+m) = (o_{T-}(n))^{-1}\left(\mathbf{I} - (o_{T-}(n))^{-1}\,\Sigma_{T-}(n,m)\Lambda_{T-}\right)^{-1}\Lambda_{T-} = (o_{T-}(n+m))^{-1}\,\Lambda_{T-}, \tag{139}$$

and the boundary condition in (115) therefore satisfies the functional equation in (116). Second, substituting (116) into the differential equation in (114), I obtain

$$0 = -C_t(n,m)\dot{M}_t(n+m)C_t(n,m) + C_t(n,m)M_t(n+m)\dot{\Sigma}_t(n,m)M_t(n+m)C_t(n,m) - \frac{A_{Q,t}^\top A_{Q,t}}{B_{Q,t}^2} \tag{140}$$

$$+ C_t(n,m)M_t(n+m)\left(\frac{B_{\Psi,t}(n)A_{Q,t}}{B_{Q,t}} - A_{\Psi,t}\right) + \left(\frac{B_{\Psi,t}(n)A_{Q,t}}{B_{Q,t}} - A_{\Psi,t}\right)^\top M_t(n+m)C_t(n,m) \tag{141}$$

where the first term on the right-hand side follows from regrouping and applying Woodbury matrix identity in (135). Multiplying both sides of this expression by $C_t(n,m)^{-1}$, regrouping and simplifying, I obtain

$$\dot{M}_t(n+m) = M_t(n+m)\left(\left(B_{\Psi,t}(n) - \frac{\Sigma_t(n,m)A_{Q,t}^\top}{B_{Q,t}}\right)\frac{A_{Q,t}}{B_{Q,t}} - A_{\Psi,t}\right)(\mathbf{I} + M_t(n+m)\Sigma_t(n,m)) \tag{142}$$

$$+ (\mathbf{I} + M_t(n+m)\Sigma_t(n,m))\left(\left(B_{\Psi,t}(n) - \frac{\Sigma_t(n,m)A_{Q,t}^\top}{B_{Q,t}}\right)\frac{A_{Q,t}}{B_{Q,t}} - A_{\Psi,t}\right)^\top M_t(n+m) \tag{143}$$

$$- \frac{A_{Q,t}^\top A_{Q,t}}{B_{Q,t}^2} + M_t(n+m)\left(\dot{\Sigma}_t(n,m) + \Sigma_t(n,m)\frac{A_{Q,t}^\top A_{Q,t}}{B_{Q,t}^2}\Sigma_t(n,m)\right)M_t(n+m). \tag{144}$$

Furthermore, notice that

$$\Sigma_t(n,m)A_{Q,t}^\top = \begin{bmatrix} 1 & -\frac{\lambda_{1,t}}{\lambda_{2,t}} \end{bmatrix}^\top k_t B_{Q,t}\frac{m}{\sigma_S^2}o_t(n)o_t(n+m) \equiv \omega_t k_t B_{Q,t}\frac{m}{\sigma_S^2}o_t(n)o_t(n+m) \tag{145}$$

which follows from regrouping and using the definitions in (61) and (86). This in turn implies that

$$B_{\Psi,t}(n) - \frac{\Sigma_t(n,m)A_{Q,t}^\top}{B_{Q,t}} = \begin{bmatrix} (o_t(n) - \frac{m}{\sigma_S^2}o_t(n)o_t(n+m) - o_t^c)k_t \\ \sigma_\Theta - \frac{\lambda_{1,t}}{\lambda_{2,t}}\left(o_t(n) - \frac{m}{\sigma_S^2}o_t(n)o_t(n+m)\right)k_t \end{bmatrix} = B_{\Psi,t}(n+m) \quad (146)$$

where the last equality follows from Lemma A.2 that $o_t(n+m) = o_t(n) - \frac{m}{\sigma_S^2}o_t(n)o_t(n+m)$. Plugging this expression back into (142) and further regrouping yields

$$\dot{M}_t(n+m) = -\frac{A_{Q,t}^\top A_{Q,t}}{B_{Q,t}^2} + M_t(n+m)\left(\frac{B_{\Psi,t}(n+m)A_{Q,t}}{B_{Q,t}} - A_{\Psi,t}\right) + \left(\frac{B_{\Psi,t}(n+m)A_{Q,t}}{B_{Q,t}} - A_{\Psi,t}\right)^\top M_t(n+m) \quad (147)$$

$$+ M_t(n+m)\begin{pmatrix} \dot{\Sigma}_t(n,m) + \Sigma_t(n,m)\left(\frac{B_{\Psi,t}(n+m)A_{Q,t}}{B_{Q,t}} - A_{\Psi,t}\right)^\top \\ +\Sigma_t(n,m)\frac{A_{Q,t}^\top A_{Q,t}}{B_{Q,t}^2}\Sigma_t(n,m) + \left(\frac{B_{\Psi,t}(n+m)A_{Q,t}}{B_{Q,t}} - A_{\Psi,t}\right)\Sigma_t(n,m) \end{pmatrix} M_t(n+m). \quad (148)$$

To further simplify (147), I use the result of Lemma B.2.

Lemma B.2. *The* $(2\times 2)-$*symmetric matrix* Σ*, as defined in* (118)*, satisfies the differential equation*

$$\dot{\Sigma}_t(n,m) = -\Sigma_t(n,m)\frac{A_{Q,t}^\top A_{Q,t}}{B_{Q,t}^2}\Sigma_t(n,m) - \Sigma_t(n,m)\left(\frac{B_{\Psi,t}(n+m)A_{Q,t}}{B_{Q,t}} - A_{\Psi,t}\right)^\top \quad (149)$$

$$-\left(\frac{B_{\Psi,t}(n+m)A_{Q,t}}{B_{Q,t}} - A_{\Psi,t}\right)\Sigma_t(n,m). \quad (150)$$

Proof. Consider first the right-hand side of (149). From (145), I obtain

$$\Sigma_t(n,m)\frac{A_{Q,t}^\top A_{Q,t}}{B_{Q,t}^2}\Sigma_t(n,m) = \omega_t^\top\omega_t\left(k_t\frac{m}{\sigma_S^2}o_t(n)o_t(n+m)\right)^2 = \Omega_t\left(k_t\frac{m}{\sigma_S^2}o_t(n)o_t(n+m)\right)^2. \quad (151)$$

Furthermore, from (146), I can write

$$\left(\frac{B_{\Psi,t}(n+m)A_{Q,t}}{B_{Q,t}} - A_{\Psi,t}\right)\Sigma_t(n,m) = \left(\frac{B_{\Psi,t}(n)A_{Q,t}}{B_{Q,t}} - A_{\Psi,t}\right)\Sigma_t(n,m) - \Omega_t\left(k_t\frac{m}{\sigma_S^2}o_t(n)o_t(n+m)\right)^2, \quad (152)$$

which, substituted in (151) yields

$$\dot{\Sigma}_t(n,m) = \Omega_t\left(k_t\frac{m}{\sigma_S^2}o_t(n)o_t(n+m)\right)^2 - \Sigma_t(n,m)\left(\frac{B_{\Psi,t}(n)A_{Q,t}}{B_{Q,t}} - A_{\Psi,t}\right)^\top - \left(\frac{B_{\Psi,t}(n)A_{Q,t}}{B_{Q,t}} - A_{\Psi,t}\right)\Sigma_t(n,m) \quad (153)$$

$$= \Omega_t\left(k_t\frac{m}{\sigma_S^2}o_t(n)o_t(n+m)\right)^2 + \left(\dot{\Omega}_t - 2\Omega_t k_t^2 o_t(n)\right)\left(\frac{m}{\sigma_S^2}o_t(n)o_t(n+m)\right) \quad (154)$$

where the second equality follows from simplifications based on (61). Second, consider the left-hand side of (149) and directly differentiate Σ in (118) using (75) to obtain

$$\dot{\Sigma}_t(n,m) = \dot{\Omega}_t \frac{m}{\sigma_S^2} o_t(n) o_t(n+m) - \Omega_t \frac{m}{\sigma_S^2} k_t^2 o_t(n) o_t(n+m)(o_t(n) + o_t(n+m)). \tag{155}$$

Finally, regrouping terms in (153) and using Lemma A.2 shows that (153) and (155) coincide and the differential equation in (149) must therefore hold. □

Substituting the differential equation of Lemma B.2 into (147) shows that the differential equation in (114) satisfies the functional equation in (116), as desired. Finally, consider the differential equation in (112). Clearly,

$$R_{T-}(n+m) = C_{T-}(n,m)^{-1}\mathbf{0} = \mathbf{0} \tag{156}$$

and the boundary condition in (112) therefore satisfies the functional equation in (113). Furthermore, substituting (113) into the differential equation in (112), regrouping and simplifying, I obtain

$$\dot{R}_t(n+m) = \begin{pmatrix} \dot{M}_t(n+m)\Sigma_t(n,m) + M_t(n+m)\dot{\Sigma}_t(n,m) \\ +C_t(n,m)^{-1}\left(\frac{B_{\Psi,t}(n)A_{Q,t}}{B_{Q,t}} - A_{\Psi,t}\right)^\top \end{pmatrix} C_t(n,m) R_t(n+m). \tag{157}$$

Using (114) and Lemma B.2, it follows that

$$\dot{M}_t(n+m)\Sigma_t(n,m) + M_t(n+m)\dot{\Sigma}_t(n,m) = -C_t(n,m)^{-1}\frac{A_{Q,t}^\top A_{Q,t}}{B_{Q,t}^2}\Sigma_t(n,m), \tag{158}$$

which, substituted into (157) yields

$$\dot{R}_t(n+m) = \left(\left(B_{\Psi,t}(n) - \frac{\Sigma_t(n,m)A_{Q,t}^\top}{B_{Q,t}}\right)\frac{A_{Q,t}}{B_{Q,t}} - A_{\Psi,t}\right)^\top R_t(n+m). \tag{159}$$

Substituting the relation in (146) into this expression shows that the differential equation in (114) satisfies the functional equation in (116), which concludes the proof. □

Theorem B.1 shows that an agent's value function, as given in (109), remains affine quadratic under continuous-discrete learning. The coefficients of the value function, which solve the system of functional and differential equations described in Theorem B.1, have explicit solutions highlighted in Lemma B.3.

Lemma B.3. *The* $(2\times 1)-$*vector coefficient* $R_t(n)$ *and the* $(2\times 2)-$*matrix coefficient* $M_t(n)$*, which satisfy the system of equations described in Theorem B.1, have explicit solutions of the form*

$$R_t(n) = \mathbf{0}, \quad \forall n \in \mathbb{N} \tag{160}$$

and

$$M_t(n) = (o_t(n))^{-1}\begin{bmatrix} (1-\lambda_{1,t})^2 & -\lambda_{2,t}(1-\lambda_{1,t}) \\ -\lambda_{2,t}(1-\lambda_{1,t}) & \lambda_{2,t}^2 \end{bmatrix} \equiv (o_t(n))^{-1}\Lambda_t, \quad \forall n \in \mathbb{N} \tag{161}$$

for any $t \in [0,T)$. Moreover, the scalar coefficient $u_t(n)$ satisfies the Boltzmann equation

$$\frac{d}{dt}u_t(n) = -\frac{1}{2}\frac{1}{o_t(n)}(k_t o_t(n) - B_{Q,t})^2 - \eta_t(n) \tag{162}$$

$$+ \eta_t(n)\sum_{m\in\mathbb{N}}\pi_t(m;n)\left(\frac{o_t(n+m)}{o_t(n)}\right)^{\frac{1}{2}}\exp\left(-(u_t(n+m) - u_t(n)\right), \quad \forall n \in \mathbb{N} \tag{163}$$

with boundary condition $u_{T-}(n) = 0$.

Proof. Consider first the functional equation in (116) and reorganize it to obtain

$$M_t(n+m) = M_t(n)(\mathbf{I} - \Sigma_t(n,m)M_t(n))^{-1}. \tag{164}$$

Evaluating this expression at $n \equiv 0$ and $m \equiv n$, I obtain

$$M_t(n) = M_t(0)(\mathbf{I} - \Sigma_t(0,n)M_t(0))^{-1} \equiv M_t^c(\mathbf{I} - \Sigma_t(0,n)M_t^c)^{-1}. \tag{165}$$

Furthermore, using the differential equation in (114) and letting $B_{\Psi,t}(0) \equiv B_{\Psi,t}$, the coefficient M_t^c satisfies

$$\dot{M}_t^c = -\frac{A_{Q,t}^\top A_{Q,t}}{B_{Q,t}^2} + M_t^c\left(\frac{B_{\Psi,t}A_{Q,t}}{B_{Q,t}} - A_{\Psi,t}\right) + \left(\frac{B_{\Psi,t}A_{Q,t}}{B_{Q,t}} - A_{\Psi,t}\right)^\top M_t^c \tag{166}$$

with boundary condition $M_{T-}^c = \frac{1}{o_{T-}^c}\Lambda_{T-}$. Conjecture that M_t^c has a solution of the form $M_t^c = m_t\Lambda_t$ where m_t is a scalar coefficient. Substituting this conjecture into (166) and multiplying both sides twice by Ω_t, I obtain

$$\Omega_t\Lambda_t\Omega_t\frac{\mathrm{d}}{\mathrm{d}t}m_t = -\Omega_t\frac{A_{Q,t}^\top A_{Q,t}}{B_{Q,t}^2}\Omega_t + \Omega_t\left(\Lambda_t\left(\frac{B_{\Psi,t}A_{Q,t}}{B_{Q,t}} - A_{\Psi,t}\right) + \left(\frac{B_{\Psi,t}A_{Q,t}}{B_{Q,t}} - A_{\Psi,t}\right)^\top\Lambda_t - \dot{\Lambda}_t\right)\Omega_t m_t. \tag{167}$$

Simplifications based on (61) then show that

$$\Omega_t\frac{A_{Q,t}^\top A_{Q,t}}{B_{Q,t}^2}\Omega_t = \Omega_t k_t^2 \tag{168}$$

and

$$\Omega_t\left(\Lambda_t\left(\frac{B_{\Psi,t}A_{Q,t}}{B_{Q,t}} - A_{\Psi,t}\right) + \left(\frac{B_{\Psi,t}A_{Q,t}}{B_{Q,t}} - A_{\Psi,t}\right)^\top\Lambda_t - \dot{\Lambda}_t\right)\Omega_t = 2k_t^2 o_t^c\Omega_t. \tag{169}$$

Observing that $\Omega_t\Lambda_t\Omega_t = \Omega_t$ and substituting this relation along with (168) and (169) into (167), the matrix differential equation in (167) reduces to an ordinary differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t}m_t = (2o_t^c m_t - 1)k_t^2, \quad m_{T-} = \frac{1}{o_{T-}^c} \tag{170}$$

where the boundary condition follows from direct comparison of the conjecture with (115). Further conjecturing that $m_t \equiv \frac{1}{o_t^c}$ and substituting this conjecture into (170) shows that

$$\frac{1}{o_t^c} = \frac{1}{o_0^c} + \int_0^t k_u^2 \mathrm{d}u, \tag{171}$$

which coincides with (82), thereby validating the conjecture. Substituting the solution $M_t^c = \frac{1}{o_t^c}\Lambda_t$ into (165) and simplifying, I obtain

$$M_t(n) = \frac{1}{o_t^c}\Lambda_t \left(\mathbf{I} - \Sigma_t(0,n)\frac{1}{o_t^c}\Lambda_t\right)^{-1} = \left(\frac{1}{o_t^c} + \frac{n}{\sigma_S^2}\right)\Lambda_t. \tag{172}$$

Second, consider the functional equation in (113) and reorganize it to obtain

$$R_t(n) = (\mathbf{I} + M_t(n)\Sigma_t(0,n))R_t^c \tag{173}$$

where $R_t^c \equiv R_t(0)$. Using the differential equation in (112), the coefficient R_t^c satisfies

$$\dot{R}_t^c = \left(\frac{B_{\Psi,t}A_{Q,t}}{B_{Q,t}} - A_{\Psi,t}\right)^\top R_t^c \tag{174}$$

with boundary condition $R_{T-}^c = 0$. Conjecture that R_t^c has a solution of the form $R_t^c = r_t\lambda_t$ where λ is the $(2\times 1)-$vector defined in Lemma A.3. Substituting this conjecture into (174) and multiplying both sides by Ω_t, I obtain

$$\omega_t \frac{\mathrm{d}}{\mathrm{d}t} r_t = \Omega_t \left(\left(\frac{B_{\Psi,t}A_{Q,t}}{B_{Q,t}} - A_{\Psi,t}\right)\lambda_t - \dot{\lambda}_t\right) r_t \tag{175}$$

where $\Omega_t\lambda_t = \omega_t$ is the $(2\times 1)-$vector defined in Lemma A.2. Simplifications based on (61) then show that

$$\Omega_t \left(\left(\frac{B_{\Psi,t}A_{Q,t}}{B_{Q,t}} - A_{\Psi,t}\right)\lambda_t - \dot{\lambda}_t\right) = k_t^2 o_t^c \omega_t. \tag{176}$$

Substituting (176) into (175), the differential equation in (175) reduces to an ordinary differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t} r_t = o_t^c k_t^2 r_t, \quad r_{T-} = 0. \tag{177}$$

The solution to the ODE in (177) is

$$r_t = r_{T-} \exp\left(-\int_t^T o_u^c k_u^2 \mathrm{d}u\right) \equiv 0 \tag{178}$$

where the last equality follows from imposing the boundary condition in (177). Finally, (162) follows from substituting (160) and (161) into (110) and simplifying. □

Substituting the coefficients of Lemma B.3 along with (109) into (106), I obtain agent i's optimal portfolio policy explicitly

$$\theta_t(\Psi^i, n^i) = \frac{A_{Q,t} - B_{Q,t}\left(o_t(n^i)\right)^{-1} B_{\Psi,t}(n^i)^\top \Lambda_t}{\gamma B_{Q,t}^2}\Psi^i. \tag{179}$$

Appendix C. Proof of Proposition 3 (Equilibrium)

In this appendix, I aggregate the optimal portfolio demands of Appendix B and impose that markets clear. Aggregation is performed based on the population distribution derived in Appendix F and the learning dynamics derived in Appendix A. I show that the equilibrium price coefficients are available in closed form.

In equilibrium, the optimal portfolio policies in (179) must satisfy the market-clearing condition

$$\int_0^1 \theta_t(\Psi^i, n^i)\mathrm{d}i = \Theta_t, \quad \forall t \in [0,T). \tag{180}$$

Imposing (180) yields a system of differential equations for the price coefficients, which I provide in Theorem C.1.

Theorem C.1. *In a linear equilibrium (see Equation 9), the price coefficients λ_1 and λ_2 satisfy the system of differential equations*

$$\sum_{n\in\mathbb{N}} \mu_t(n)\left(A_{Q,t} - B_{Q,t}\left(o_t(n)\right)^{-1} B_{\Psi,t}(n)^\top \Lambda_t\right)\Gamma_t(n) = \mathbf{1}^\star \gamma B_{Q,t}^2 \tag{181}$$

with boundary condition

$$\begin{bmatrix} \lambda_{1,T-} & \lambda_{2,T-} \end{bmatrix} = \begin{bmatrix} \frac{\phi_T}{\sigma_S^2} & -\gamma \end{bmatrix} \frac{o_{T-}^c \sigma_S^2}{\phi_T o_{T-}^c + \sigma_S^2} \tag{182}$$

where ϕ denotes the cross-sectional average number of signals, $\mathbf{1}^\star$ is a $(1\times 2)-$vector defined as $\mathbf{1}^\star = \begin{bmatrix} 0 & 1 \end{bmatrix}$ and Γ is a $(2\times 2)-$matrix defined as

$$\Gamma_t(n) = \begin{bmatrix} \begin{matrix} \alpha_t(n) \\ \frac{\lambda_{1,t}}{\lambda_{2,t}}(1-\alpha_t(n)) \end{matrix} & (\mathbf{1}^\star)^\top \end{bmatrix} \tag{183}$$

with $\alpha_t(n) \equiv \frac{o_t^c n}{\sigma_S^2 + o_t^c n}$.

Proof. I start by providing an aggregation result regarding the conditional expectation Ψ^i, which I then use to aggregate optimal demands. Let $\widehat{\Psi}_t \equiv \int_0^1 \Psi_t^i \mathrm{d}i$ denote the market average expectation of Ψ_t. Lemma C.1 then provides the relation between $\widehat{\Psi}_t$ and Ψ_t.

Lemma C.1. *The average market expectation $\widehat{\Psi}_t$ at time $t \in [0,T)$ satisfies*

$$\widehat{\Psi}_t = \sum_{n\in\mathbb{N}} \mu_t(n)\Gamma_t(n)\Psi_t \tag{184}$$

where the $(2\times 2)-$matrix Γ is defined in (183).

Proof. Using (75), (79) and (98), an application of Ito's lemma yields

$$\mathrm{d}\frac{\widehat{\Pi}_t^c}{o_t^c} = k_t^2 \widehat{\Pi}_t^i \mathrm{d}t + k_t \mathrm{d}\widehat{B}_t^i \tag{185}$$

and

$$\mathrm{d}\frac{\widehat{\Pi}_t^i}{o_t^i} = k_t^2 \widehat{\Pi}_{t-}^i \mathrm{d}t + k_t \mathrm{d}\widehat{B}_t^i + \left(\frac{\widehat{\Pi}_t^i}{o_t^i} - \frac{\widehat{\Pi}_{t-}^i}{o_{t-}^i}\right) \mathrm{d}N_t^i \tag{186}$$

$$= k_t^2 \widehat{\Pi}_{t-}^i \mathrm{d}t + k_t \mathrm{d}\widehat{B}_t^i + \frac{\Delta n_t^i}{\sigma_S^2} Y_t^i \mathrm{d}N_t^i \tag{187}$$

where the second equality in (186) follows from

$$\widehat{\Pi}_t^i = \widehat{\Pi}_{t-}^i + o_t^i \frac{\Delta n_t^i}{\sigma_S^2} \widehat{Y}_t^i = \frac{o_t^i}{o_{t-}^i} \widehat{\Pi}_{t-}^i + o_t^i \frac{\Delta n_t^i}{\sigma_S^2} Y_t^i. \tag{188}$$

Subtracting (185) from (186) and integrating, I obtain

$$\frac{\widehat{\Pi}_t^i}{o_t(n_t^i)} = \frac{\widehat{\Pi}_t^c}{o_t^c} + \sum_{s \le t} \frac{\Delta n_s^i}{\sigma_S^2} Y_s^i \Delta N_s^i. \tag{189}$$

Using Lemma A.1 and reorganizing, this expression becomes

$$\widehat{\Pi}_t^i = \frac{o_t(n_t^i)}{o_t^c} \widehat{\Pi}_t^c + \frac{o_t(n_t^i) n_t^i}{\sigma_S^2} \Pi + \frac{o_t(n_t^i)}{\sigma_S} \sum_{s \le t} \left(\Delta n_s^i\right)^{\frac{1}{2}} \epsilon_s^i \Delta N_s^i. \tag{190}$$

Letting $\widehat{\Pi}_t \equiv \int_0^1 \widehat{\Pi}_t^i \mathrm{d}i$ denote the market average expectation of Π and using (81) and the law of large numbers whereby $\int_0^1 \epsilon_s^i \mathrm{d}i = 0$, $\forall s \in (0, T)$, I obtain

$$\widehat{\Pi}_t = \sum_{n \in \mathbb{N}} \mu_t(n) \left(\frac{\sigma_S^2}{\sigma_S^2 + o_t^c n} \widehat{\Pi}_t^c + \frac{o_t^c n}{\sigma_S^2 + o_t^c n} \Pi \right) \tag{191}$$

$$\equiv \sum_{n \in \mathbb{N}} \mu_t(n) \Big((1 - \alpha_t(n)) \widehat{\Pi}_t^c + \alpha_t(n) \Pi \Big). \tag{192}$$

The relation in (184) then follows from vectorizing

$$\int_0^1 \Delta_t^i \mathrm{d}i = \widehat{\Pi}_t - \widehat{\Pi}_t^c = \sum_{n \in \mathbb{N}} \mu_t(n) \alpha_t(n) \Delta_t \tag{193}$$

and

$$\int_0^1 \widehat{\Theta}_t^i \mathrm{d}i = \frac{\lambda_{1,t}}{\lambda_{2,t}} (\Pi - \widehat{\Pi}_t) + \Theta_t = \frac{\lambda_{1,t}}{\lambda_{2,t}} \sum_{n \in \mathbb{N}} \mu_t(n) (1 - \alpha_t(n)) \Delta_t + \Theta_t. \tag{194}$$

□

Remark 3. *As in He and Wang (1995), the relation in* (191) *allows me to define a relation between* *first-order and higher-order expectations (HOE). Letting* $\alpha_t \equiv \sum_{n\in\mathbb{N}} \mu_t(n)\alpha_t(n)$*, notice that agent* i*'s expectation of market average expectations* $\widehat{\Pi}$ *is*

$$\mathbb{E}\left[\widehat{\Pi}_t \middle| \mathscr{F}_t^i\right] = (1-\alpha_t)\widehat{\Pi}_t^c + \alpha_t \widehat{\Pi}_t^i. \tag{195}$$

Hence, the second-order expectation $\widehat{\Pi}_t^{(2)}$ *of* Π *satisfies*

$$\widehat{\Pi}_t^{(2)} = \int_0^1 \mathbb{E}\left[\widehat{\Pi}_t \middle| \mathscr{F}_t^i\right] di = (1-\alpha_t^2)\widehat{\Pi}_t^c + \alpha_t^2 \Pi = (1-\alpha_t)\widehat{\Pi}_t^c + \alpha_t \widehat{\Pi}_t. \tag{196}$$

Iterating, the $k-$*th order expectation* $\widehat{\Pi}^{(k)}$ *satisfies the recursion*

$$\widehat{\Pi}_t^{(k)} = \int_0^1 \mathbb{E}\left[\widehat{\Pi}_t^{(k-1)} \middle| \mathscr{F}_t^i\right] di = (1-\alpha_t)\widehat{\Pi}_t^c + \alpha_t \widehat{\Pi}_t^{(k-1)} = (1-\alpha_t^k)\widehat{\Pi}_t^c + \alpha_t^k \Pi. \tag{197}$$

HOE are a weighted average of the first-order expectation $\widehat{\Pi}$ *and common expectations* $\widehat{\Pi}^c$*.*

To obtain the boundary condition in (182), I substitute the optimal demands in (122) into (180). Using the aggregation result of Lemma C.1, I obtain

$$\int_0^1 \theta_{T-}^i \mathrm{d}i = \sum_{n\in\mathbb{N}} \mu_T(n) \frac{1}{\gamma} \left(o_{T-}(n)\right)^{-1} \lambda_{T-}^\top \Gamma_{T-}(n)\Psi_T = \mathbf{1}^\star \Psi_T. \tag{198}$$

Using $\mathbf{L}(\cdot)$ to denote a general linear relation, an important observation is then that the term inside the sum is linear in n:

$$\left(o_{T-}(n)\right)^{-1} \lambda_{T-}^\top \Gamma_{T-}(n) \equiv \mathbf{L}(n) = \begin{bmatrix} \frac{n}{\sigma_S^2} - \frac{n o_{T-}^c + \sigma_S^2}{o_{T-}^c \sigma_S^2}\lambda_{1,T-} & -\frac{n o_{T-}^c + \sigma_S^2}{o_{T-}^c \sigma_S^2}\lambda_{2,T-} \end{bmatrix}. \tag{199}$$

As a result, the summation only involves the cross-sectional average number of signals ϕ_T at time T

$$\begin{bmatrix} \frac{\phi_T}{\sigma_S^2} - \frac{\phi_T o_{T-}^c + \sigma_S^2}{o_{T-}^c \sigma_S^2}\lambda_{1,T-} & -\frac{\phi_T o_{T-}^c + \sigma_S^2}{o_{T-}^c \sigma_S^2}\lambda_{2,T-} \end{bmatrix} \Psi_T = \gamma \mathbf{1}^\star \Psi_T. \tag{200}$$

Separating variables and regrouping terms then yields (182). Finally, to obtain the system of differentials equations in (181), I substitute the optimal demands in (179) into (180). Using Lemma C.1, I obtain

$$\frac{1}{\gamma B_{Q,t}^2} \sum_{n\in\mathbb{N}} \mu_t(n) \left(A_{Q,t} - B_{Q,t}\left(o_t(n)\right)^{-1} B_{\Psi,t}(n)^\top \Lambda_t\right) \Gamma_t(n)\Psi_t = \mathbf{1}^\star \Psi_t \tag{201}$$

Separating variables directly yields (179). □

Remark 4. *Equations* (200) *and* (201) *imply that* $\mathbf{L}(P_t, \Psi_t) = 0$ *and, as a result, the price satisfies* $P_t = \mathbf{L}(\Psi_t)$*,* $\forall t \in [0, T)$*. Hence, the equilibrium price satisfies the linear conjecture in Equation 9. The specific functional form of prices in Equation 9 follows from* (191)*.*

From the system of differential equations in (181), I obtain a closed-form solution for the ratio of the price coefficients λ_1 and λ_2, which I formulate in Lemma C.2.

Lemma C.2. *In equilibrium, the signal-noise ratio of the price admits the closed-form solution*

$$\frac{\lambda_{1,t}}{\lambda_{2,t}} = -\frac{\phi_t}{\gamma\sigma_S^2}, \quad \forall t \in [0, T). \tag{202}$$

It follows that common uncertainty o^c also admits a closed-form solution, which satisfies

$$o_t^c = \left(\frac{1}{\sigma_\Pi^2} + \left(\frac{\phi_0}{\gamma\sigma_\Theta\sigma_S^2} \right)^2 + \left(\frac{1}{\gamma\sigma_\Theta\sigma_S^2} \right)^2 \int_0^t \left(\frac{d}{ds}\phi_s + a_\Theta\phi_s \right)^2 dt \right)^{-1}, \quad \forall t \in [0, T). \tag{203}$$

Proof. Multiplying both sides of (181) by ω yields

$$\sum_{n\in\mathbb{N}} \mu_t(n) \left(A_{Q,t} - B_{Q,t} \left(o_t(n)\right)^{-1} B_{\Psi,t}(n)^\top \Lambda_t \right) \Gamma_t(n)\omega_t = -\frac{\lambda_{1,t}}{\lambda_{2,t}} \gamma B_{Q,t}^2 \tag{204}$$

Using (61), notice that

$$A_{Q,t}\Gamma_t(n)\omega_t = \frac{no_t^c}{no_t^c + \sigma_S^2} k_t B_{Q,t} \tag{205}$$

and

$$(o_t(n))^{-1} B_{\Psi,t}(n)^\top \Lambda_t \Gamma_t(n)\omega_t = \frac{no_t^c}{no_t^c + \sigma_S^2} k_t - B_{Q,t}\frac{n}{\sigma_S^2}. \tag{206}$$

Plugging these expressions into (204), I obtain an algebraic equation

$$\sum_{n\in\mathbb{N}} \mu_t(n) \frac{n}{\sigma_S^2} B_{Q,t}^2 = -\frac{\lambda_{1,t}}{\lambda_{2,t}} \gamma B_{Q,t}^2. \tag{207}$$

Observing that $\frac{n}{\sigma_S^2} B_{Q,t}^2 = \mathbf{L}(n)$, the summation only involves the cross-sectional average number of signals ϕ and the equation becomes

$$\frac{\phi_t}{\sigma_S^2} = -\frac{\lambda_{1,t}}{\lambda_{2,t}} \gamma, \quad \forall t \in [0, T) \tag{208}$$

from which I obtain (202). Inspection of the boundary condition in (182) reveals that it also satisfies (202). Equation (203) directly follows from substituting (202) into (82). □

Furthermore, assuming that the noisy supply follows a random walk and using the result of Lemma C.2, the price coefficient λ_2 has a closed-form solution, which I highlight in Lemma C.3.

Lemma C.3. *Assuming that noise-trading demand follows a martingale (i.e., $a_\Theta \equiv 0$), the price coefficient λ_2 admits the closed-form solution*

$$\lambda_{2,t} = -\gamma \frac{\sigma_S^2 o_t^c}{\sigma_S^2 + \phi_t o_t^c}, \quad \forall t \in [0, T). \tag{209}$$

Proof. Spelling out the system of equations in (181) and simplifying, the second differential equation is

$$\frac{\mathrm{d}}{\mathrm{d}t}\lambda_{2,t} = a_\Theta \lambda_{2,t} + \frac{B_{Q,t}}{o_t^c \sigma_S^2}\lambda_{2,t}\left(B_{Q,t}\sum_{n\in\mathbb{N}}\mu_t(n)(no_t^c + \sigma_S^2) - \sigma_S^2 o_t^c k_t\right) + \gamma B_{Q,t}^2. \tag{210}$$

Observing that $no_t^c + \sigma_S^2 = \mathbf{L}(n)$, the summation only involves the cross-sectional average number of signals ϕ and the equation becomes

$$\frac{\mathrm{d}}{\mathrm{d}t}\lambda_{2,t} = a_\Theta \lambda_{2,t} - k_t \lambda_{2,t} B_{Q,t} + B_{Q,t}^2\left(\lambda_{2,t}\frac{\phi_t o_t^c + \sigma_S^2}{o_t^c \sigma_S^2} + \gamma\right). \tag{211}$$

Furthermore, assuming that $a_\Theta \equiv 0$ and substituting (209) into (211), I get

$$\frac{\mathrm{d}}{\mathrm{d}t}\lambda_{2,t} = -k_t \lambda_{2,t} B_{Q,t}. \tag{212}$$

To finally show that this equation holds, substitute first (209) in the left-hand side of (212) and obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\lambda_{2,t} = -\gamma \frac{\sigma_S^2\left(\sigma_S^2 \frac{\mathrm{d}}{\mathrm{d}t}o_t^c - (o_t^c)^2 \frac{\mathrm{d}}{\mathrm{d}t}\phi_t\right)}{(\sigma_S^2 + \phi_t o_t^c)^2} \tag{213}$$

$$= \frac{(o_t^c)^2 \frac{\mathrm{d}}{\mathrm{d}t}\phi_t\left(\gamma^2 \sigma_S^2 \sigma_\Theta^2 + \frac{\mathrm{d}}{\mathrm{d}t}\phi_t\right)}{\gamma \sigma_\Theta^2 (\sigma_S^2 + \phi_t o_t^c)^2} \tag{214}$$

where the second line follows from substituting (202) and using (79). Substituting (209) in the right-hand side of (212) and using (202), I further obtain

$$-k_t \lambda_{2,t} B_{Q,t} = \frac{(o_t^c)^2 \frac{\mathrm{d}}{\mathrm{d}t}\phi_t\left(\gamma^2 \sigma_S^2 \sigma_\Theta^2 + \frac{\mathrm{d}}{\mathrm{d}t}\phi_t\right)}{\gamma \sigma_\Theta^2 (\sigma_S^2 + \phi_t o_t^c)^2}. \tag{215}$$

Clearly, (213) and (215) coincide and (212) must therefore hold. Finally, inspecting the boundary condition in (182) shows that it also satisfies (209), which concludes the proof. □

Appendix D. Proof of Proposition 4 (Holdings)

In this appendix, I derive an explicit expression for the informational holding (Definition 1). Using (179), I can write the portfolio θ^i of manager i as

$$\theta_t^i = d_{\Delta,t}(n^i)\Delta_t^i + d_{\Theta,t}(n^i)\widehat{\Theta}_t^i \tag{216}$$

$$= d_{\Theta,t}(n^i)\Theta_t + d_{\Delta,t}(n^i)\Delta_t^i + \frac{\lambda_{1,t}}{\lambda_{2,t}} d_{\Theta,t}(n^i)(\Pi - \widehat{\Pi}_t^i) \tag{217}$$

where, using Lemma C.2 and C.3 and Equation (61), the coefficients d_Δ and d_Θ satisfy

$$d_{\Theta,t}(n) = \frac{\sigma_S^2 + no_t^c}{\sigma_S^2 + o_t^c \phi_t}, \tag{218}$$

$$d_{\Delta,t}(n) = \frac{1}{\gamma o_t^c} d_{\Theta,t}(n). \tag{219}$$

Furthermore, using Definition 1 and Equation (216), I can write the informational holding $\widehat{\theta}^i$ of manager i as

$$\widehat{\theta}_t^i = d_{\Delta,t}(n^i)\Delta_t^i + \frac{\lambda_{1,t}}{\lambda_{2,t}} d_{\Theta,t}(n^i)(\Pi - \widehat{\Pi}_t^i) + \left(d_{\Theta,t}(n^i) - 1\right)\Theta_t. \tag{220}$$

Substituting the expression for $\widehat{\Pi}^i$ in (190) and reorganizing, I obtain

$$\widehat{\theta}_t^i = d_{\Delta,t}(n)\left(\alpha_t(n)(\Pi - \widehat{\Pi}_t^c) + \frac{o_t(n)}{\sigma_S}\sqrt{n}\epsilon_t^i\right) + \left(d_{\Theta,t}(n^i) - 1\right)\Theta_t \tag{221}$$

$$+ \frac{\lambda_{1,t}}{\lambda_{2,t}} d_{\Theta,t}(n)\left((1 - \alpha_t(n))(\Pi - \widehat{\Pi}_t^c) - \frac{o_t(n)}{\sigma_S}\sqrt{n}\epsilon_t^i\right) \tag{222}$$

$$\equiv \left(d_{\Delta,t}(n)\alpha_t(n) + \frac{\lambda_{1,t}}{\lambda_{2,t}} d_{\Theta,t}(n)(1 - \alpha_t(n))\right)\Delta_t + \left(d_{\Delta,t}(n) - \frac{\lambda_{1,t}}{\lambda_{2,t}} d_{\Theta,t}(n)\right)\frac{o_t(n)}{\sigma_S}\sqrt{n}\epsilon_t^i + \frac{o_t^c(n - \phi_t)}{\sigma_S^2 + o_t^c\phi_t}\Theta_t \tag{223}$$

where I use the fact that incremental number of signals Δn^i in (190) are independent and therefore $\sum_{s\le t}\left(\Delta n_s^i\right)^{\frac{1}{2}}\epsilon_s^i \Delta N_s^i \sim \sqrt{n_t^i}\epsilon_t^i$. Finally, using Theorem C.1 and simplifying shows that

$$d_{\Delta,t}(n)\alpha_t(n) + \frac{\lambda_{1,t}}{\lambda_{2,t}} d_{\Theta,t}(n)(1 - \alpha_t(n)) = \frac{n - \phi_t}{\gamma(\sigma_S^2 + o_t^c\phi_t)} \tag{224}$$

and

$$\left(d_{\Delta,t}(n) - \frac{\lambda_{1,t}}{\lambda_{2,t}} d_{\Theta,t}(n)\right)\frac{o_t(n)}{\sigma_S} = \frac{1}{\gamma\sigma_S}, \tag{225}$$

which substituted in (221) yields the decomposition of informational holdings in (18).

Appendix E. Proof of Proposition 5 (Alpha)

In this appendix, I derive explicit expressions for the unconditional estimate of a manager's alpha, its standard error and its $t-$statistic.

Matching the performance regression in (21) to a manager's budget constraint in (5), the econometrician computes a manager i's instantaneous alpha according to:

$$\alpha_t^i = \frac{1}{\mathrm{d}t}\widehat{\theta}_t^i \mathbb{E}\left[\mathrm{d}Q_t|\mathscr{F}_t\right] = \left(\frac{n_t^i - \phi_t}{\gamma(\sigma_S^2 + o_t^c\phi_t)}(\Delta_t + \gamma o_t^c\Theta_t) + \frac{\sqrt{n_t^i}}{\gamma\sigma_S}\epsilon_t^i\right) A_{Q,t}\Psi_t \tag{226}$$

$$\equiv (a_t(n_t^i)(\Delta_t + \gamma o_t^c\Theta_t) + b(n_t^i)\epsilon_t^i)A_{Q,t}\Psi_t \tag{227}$$

where I used the informational portfolio decomposition of Proposition 4 and where $\mathscr{F}$ denotes the econometrician's filtration defined in (20). The unconditional estimator $\widehat{\alpha}^i$ of alpha for a manager i who holds n^i ideas in turn satisfies

$$\widehat{\alpha}_t(n^i) = \mathbb{E}[\alpha_t^i|n_t^i = n^i] = \frac{1}{\mathrm{d}t}\mathbb{E}\left[\widehat{\theta}_t^i \mathrm{d}Q_t\middle| n_t^i = n^i\right] = a_t(n^i)\mathbb{E}[(\Delta_t + \gamma o_t^c\Theta_t)A_{Q,t}\Psi_t] \tag{228}$$

where I used that ϵ^i is an independent Gaussian variable with mean zero. Furthermore, using Lemma C.2 and C.3 and Equation (61) and simplifying, the vector A_Q defined in Lemma A.3 is explicitly given by

$$A_{Q,t} = \frac{\sigma_S^4 k_t (k_t - \gamma\sigma_\Theta) o_t^c}{(\sigma_S^2 + o_t^c \phi_t)^2} \begin{bmatrix} 1 \\ \gamma o_t^c \end{bmatrix}^\top . \tag{229}$$

Substituting A_Q back into (228) and rearranging yields

$$\widehat{\alpha}_t(n^i) = a_t(n^i) \frac{\sigma_S^4 k_t (k_t - \gamma\sigma_\Theta) o_t^c}{(\sigma_S^2 + o_t^c \phi_t)^2} \mathbb{E}\left[(\Delta_t + \gamma o_t^c \Theta_t)^2\right], \tag{230}$$

from which the estimator in (23) follows. To obtain the unconditional expectation $\mathbb{E}\left[(\Delta_t + \gamma o_t^c \Theta_t)^2\right]$, observe that (8) with $a_\Theta \equiv 0$ yields $\Theta_t = \Theta_0 + \sigma_\Theta B_t^\Theta$. Hence, the unconditional expectation in (230) is explicitly given by

$$\begin{aligned}
\mathbb{E}\left[(\Delta_t + \gamma o_t^c \Theta_t)^2\right] &= \mathbb{E}\left[\Delta_t^2\right] + 2\gamma o_t^c \mathbb{E}\left[\Delta_t \Theta_t\right] + (\gamma o_t^c)^2 (t+1)\sigma_\Theta^2 && (231)\\
&= \mathbb{E}\left[\mathbb{E}[(\Pi - \widehat{\Pi}_t^c)^2 | \mathscr{F}_t^c]\right] + 2\gamma o_t^c \mathbb{E}\left[\mathbb{E}[(\Pi - \widehat{\Pi}_t^c)\Theta_t | \mathscr{F}_t^c]\right] + (\gamma o_t^c)^2 (t+1)\sigma_\Theta^2 && (232)\\
&= o_t^c - 2\gamma (o_t^c)^2 \frac{\lambda_{1,t}}{\lambda_{2,t}} + (\gamma o_t^c)^2 (t+1)\sigma_\Theta^2 && (233)\\
&= o_t^c + 2(o_t^c)^2 \phi_t / \sigma_S^2 + (\gamma o_t^c)^2 (t+1)\sigma_\Theta^2, && (234)
\end{aligned}$$

which yields Eq. (25).

The standard error σ^i of a manager i's instantaneous alpha in (226) is the square root of the quadratic variation of the orthogonal noise in (21), which is given by

$$\begin{aligned}
\sigma_t^i &= \sqrt{\frac{1}{\mathrm{d}t}(a_t(n_t^i)(\Delta_t + \gamma o_t^c \Theta_t) + b(n_t^i)\epsilon_t^i)^2 \mathrm{d}\langle Q\rangle_t} = |B_{Q,t}| |a_t(n_t^i)(\Delta_t + \gamma o_t^c \Theta_t) + b(n_t^i)\epsilon_t^i| && (235)\\
&= \frac{\sigma_S^2(|k_t| + \gamma\sigma_\Theta) o_t^c}{\sigma_S^2 + o_t^c \phi_t} |a_t(n_t^i)(\Delta_t + \gamma o_t^c \Theta_t) + b(n_t^i)\epsilon_t^i| && (236)
\end{aligned}$$

where the second equality follows from using Lemma C.2 and C.3 to write the diffusion B_Q (defined in Lemma A.3) explicitly as $B_{Q,t} = \frac{\sigma_S^2(k_t - \gamma\sigma_\Theta) o_t^c}{\sigma_S^2 + o_t^c \phi_t}$. Using (235), a manager i's instantaneous alpha $t-$statistic t_α^i is the ratio of alpha to its standard error under the econometrician's filtration $\mathscr{F}$:

$$t_{\alpha,t}^i = \frac{\alpha_t^i}{\sigma_t^i} = \frac{\widehat{\theta}_t^i}{\left|\widehat{\theta}_t^i\right|} \frac{A_{Q,t}\Psi_t}{|B_{Q,t}|} = \operatorname{sign}\left(\widehat{\theta}_t^i\right) \frac{\sigma_S^2 |k_t|}{\sigma_S^2 + o_t^c \phi_t} (\Delta_t + \gamma o_t^c \Theta_t) \tag{237}$$

where the second line follows from substituting (229) and the expression for B_Q above and simplifying. It follows that the unconditional $t-$statistic $\widehat{t}_\alpha^i$ of the estimator $\widehat{\alpha}^i$ for a manager who holds n^i ideas is given by

$$\widehat{t}_{\alpha,t}(n^i) = \mathbb{E}\left[t_{\alpha,t}^i \middle| n_t^i = n^i\right] = \frac{\sigma_S^2 |k_t|}{\sigma_S^2 + o_t^c \phi_t} \mathbb{E}\left[\operatorname{sign}\left(\widehat{\theta}_t^i\right)(\Delta_t + \gamma o_t^c \Theta_t) \middle| n_t^i = n^i\right]. \tag{238}$$

To compute the expectation in (238), define the variable

$$X_t := \Delta_t + \gamma o_t^c \Theta_t \sim \mathscr{N}\left(0, \mathbb{E}\left[X_t^2\right]\right) \tag{239}$$

and observe that

$$\operatorname{sign}(\widehat{\theta}_t^i) = 1 \Leftrightarrow \begin{cases} X_t > -\frac{b(n)}{a_t(n)}\epsilon_t^i, & n \geq \phi_t \\ X_t < -\frac{b(n)}{a_t(n)}\epsilon_t^i, & n < \phi_t \end{cases} \tag{240}$$

with the same result applying symmetrically for $\operatorname{sign}(\widehat{\theta}_t^i) = -1$. Using this result, I compute the expectation in (238) as

$$\mathbb{E}\left[\operatorname{sign}\left(\widehat{\theta}_t^i\right) X_t\right] = (-1)^{\mathbf{1}_{n<\phi_t}} \mathbb{E}\left[\begin{array}{c} \mathbb{P}\left[X_t > -\frac{b(n)}{a_t(n)}\epsilon\right] \mathbb{E}\left[X_t \middle| X_t > -\frac{b(n)}{a_t(n)}\epsilon\right] \\ -\mathbb{P}\left[X_t < -\frac{b(n)}{a_t(n)}\epsilon\right] \mathbb{E}\left[X_t \middle| X_t < -\frac{b(n)}{a_t(n)}\epsilon\right] \end{array}\right] \tag{241}$$

$$= (-1)^{\mathbf{1}_{n<\phi_t}} \frac{1}{\sqrt{2\pi \mathbb{E}\left[X_t^2\right]}} \mathbb{E}\left[\int_{-\frac{b(n)}{a_t(n)}\epsilon}^{\infty} X_t e^{-\frac{1}{2}\frac{X_t^2}{\mathbb{E}[X_t^2]}} \mathrm{d}X_t - \int_{-\infty}^{-\frac{b(n)}{a_t(n)}\epsilon} X_t e^{-\frac{1}{2}\frac{X_t^2}{\mathbb{E}[X_t^2]}} \mathrm{d}X_t\right] \tag{242}$$

$$= (-1)^{\mathbf{1}_{n<\phi_t}} \sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{\mathbb{E}[X_t^2]}} \mathbb{E}\left[\int_{-\frac{b(n)}{a_t(n)}\epsilon}^{\infty} X_t e^{-\frac{1}{2}\frac{X_t^2}{\mathbb{E}[X_t^2]}} \mathrm{d}X_t\right] \tag{243}$$

$$= (-1)^{\mathbf{1}_{n<\phi_t}} \sqrt{\frac{2}{\pi}} \sqrt{\mathbb{E}[X_t^2]} \int_{\mathbb{R}} \exp\left(-\frac{1}{2}\frac{1}{\mathbb{E}[X_t^2]}\left(\frac{b(n)}{a_t(n)}\right)^2 \epsilon^2\right) \phi(\epsilon)\mathrm{d}\epsilon \tag{244}$$

$$= (-1)^{\mathbf{1}_{n<\phi_t}} \sqrt{\frac{2}{\pi}} \frac{\mathbb{E}[X_t^2]}{\sqrt{\mathbb{E}[X_t^2] + \left(\frac{b(n)}{a_t(n)}\right)^2}}. \tag{245}$$

Substituting (241) in (238), I finally obtain (24).

Appendix F. Population dynamics

In this appendix, I derive the dynamics of information sharing for the general setting of Section 2.1 (Appendix F.1). I then derive an explicit solution for the average number of ideas under network formation in Section 5.2.2 (Appendix F.2). Finally, I derive a manager's average trajectory of number of ideas, which I use for the main analysis in Section 5 (Appendix F.3).

F.1. Population dynamics in the general setting of Section 2.1

In the general setting of Section 2.1, the dynamics of a manager i's number n^i of ideas satisfy

$$\mathrm{d}n_t^i = \Delta n_t^i \mathrm{d}N_t^i, \quad n_0^i \sim \pi_0, \quad \Delta n_t^i \sim \pi_t(\cdot; n_{t-}^i) \tag{246}$$

where $(N^i)_{t\geq 0}$ denotes a Poisson process with intensity $\eta(n_{t-}^i)$. These dynamics imply a certain cross-sectional distribution, μ, of number of ideas, i.e., the distribution μ must satisfy a certain Kolmogorov Forward Equation (KFE), which I derive using the result formulated in Lemma F.1.

Lemma F.1. *Define the expectation*

$$g_t \equiv \mathbb{E}[f(n_t)] = \sum_{k\in\mathbb{N}} f(k)\mu_t(k) \tag{247}$$

for some continuous function $f(\cdot)$. Then, the function g must satisfy the differential equation

$$\frac{d}{dt}g_t = \sum_{n\in\mathbb{N}} \eta_t(n)\mu_t(n) \sum_{m\in\mathbb{N}} f(n+m)\pi_t(m;n) - \sum_{n\in\mathbb{N}} \eta_t(n)\mu_t(n)f(n) \tag{248}$$

with initial condition $g_0 = \sum_{k\in\mathbb{N}} f(k)\pi_0(k)$.

Proof. Observe that the generator of the process in (246) satisfies

$$Af(n) = \eta_t(n) \sum_{m\in\mathbb{N}} \pi_t(m;n)(f(n+m) - f(n)) \tag{249}$$

and rewrite the expectation in (247) as

$$g_t = \mathbb{E}[f(n_0)] + \int_0^t \mathbb{E}[Af(n_s)]\mathrm{d}s = \mathbb{E}[f(n_0)] + \int_0^t \sum_{n\in\mathbb{N}} Af(n)\mu_s(n)\mathrm{d}s. \tag{250}$$

Differentiating Eq. (250) with respect to time and rearranging yields (248). □

To obtain the KFE in (3), I then change the summation order in (248). Specifically, introduce the change of variable $k \equiv n+m$ and rewrite the first term in (248) as

$$\sum_{n\in\mathbb{N}}\sum_{m\in\mathbb{N}} f(n+m)\eta_t(n)\pi_t(m;n)\mu_t(n) = \sum_{k\in\mathbb{N}}\sum_{m=1}^{k-1} f(k)\eta_t(k-m)\pi_t(m;k-m)\mu_t(k-m). \tag{251}$$

Plugging (251) into (248) and rearranging, I obtain

$$\sum_{n\in\mathbb{N}} f(n)\frac{\mathrm{d}}{\mathrm{d}t}\mu_t(n) = \sum_{n\in\mathbb{N}} f(n)\left(\eta\sum_{m=1}^{n-1} \eta_t(n-m)\pi_t(m;n-m)\mu_t(n-m) - \eta_t(n)\mu_t(n)\right). \tag{252}$$

Observing that this equation must hold for any continuous function $f(\cdot)$, the KFE in (3) follows.

F.2. Proof of Corollary 4

In the particular case of Section 5.2.2, I further assume that a manager belongs to two different classes, indexed by $k \in \{A, B\}$, depending on whether her number n^i of ideas is in $A = \{n \in \mathbb{N}^\star : n < N\}$ or $B = \{n \in \mathbb{N}^\star : n \geq N\}$, for some integer $N \geq 1$. Any manager is matched with some other manager with intensity η. Upon meeting someone, a manager is matched with someone of her network with probability p and with someone of the other network with probability $1-p$. Hence, two managers i and j are matched with intensity

$$\eta(n^i, n^j) = \eta\left(p(\mathbf{1}_{n^i\in A}\mathbf{1}_{n^j\in A} + \mathbf{1}_{n^i\in B}\mathbf{1}_{n^j\in B}) + (1-p)(\mathbf{1}_{n^i\in A}\mathbf{1}_{n^j\in B} + \mathbf{1}_{n^i\in B}\mathbf{1}_{n^j\in A})\right). \tag{253}$$

It follows that, conditional on having n signals at time t, manager i meets someone with intensity

$$\eta_t(n) = \frac{1}{\mathrm{d}t}\mathbb{P}\left[\mathrm{d}N_t^i = 1 \middle| n_{t-}^i = n\right] = \sum_{m\in\mathbb{N}} \eta(n,m)\mu_t(m) \tag{254}$$

$$= \eta\left(\mathbf{1}_{n\in A}\left(pq_t + (1-p)(1-q_t)\right) + \mathbf{1}_{n\in B}\left((1-p)q_t + p(1-q_t)\right)\right) \tag{255}$$

where $q_t = \sum_{m\in A}\mu_t(m)$ represents the fraction of managers in class A at time t. Upon meeting manager j at time t, manager i gets m additional signals according to

$$\pi_t(m;n) = \left(\mathbf{1}_{n\in A}\frac{p\mathbf{1}_{m\in A} + (1-p)\mathbf{1}_{m\in B}}{pq_t + (1-p)(1-q_t)} + \mathbf{1}_{n\in B}\frac{(1-p)\mathbf{1}_{m\in A} + p\mathbf{1}_{m\in B}}{p(1-q_t) + (1-p)q_t}\mu_t(m)\right)\mu_t(m). \tag{256}$$

Substituting Eqs. (254) and (256) into the KFE in Eq. (3) and simplifying yields the following population dynamics:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mu_t(n) = \eta\left(\mathbf{1}_{n\leq 2(N-1)}p\sum_{m=1\vee(n-(N-1))}^{(n-1)\wedge(N-1)}\mu_t(m)\mu_t(n-m) + \mathbf{1}_{n\geq N+1}(1-p)\sum_{m=N\vee(n-(N-1))}^{n-1}\mu_t(m)\mu_t(n-m)\right) \tag{257}$$

$$+\,\eta\left(\mathbf{1}_{n\geq N+1}(1-p)\sum_{m=1}^{(n-N)\wedge(N-1)}\mu_t(m)\mu_t(n-m) + \mathbf{1}_{n\geq 2N}p\sum_{m=N}^{n-N}\mu_t(m)\mu_t(n-m)\right) \tag{258}$$

$$-\,\eta\left(\mathbf{1}_{n\in A}\left(pq_t + (1-p)(1-q_t)\right) + \mathbf{1}_{n\in B}\left(p(1-q_t) + (1-p)q_t\right)\right)\mu_t(n), \quad \mu_0(n) = \delta_{n=1}. \tag{259}$$

To obtain the cross-sectional average number ϕ of ideas in (45), notice that the KFE in (257) decouples into a simpler equation for all $n \in A$:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mu_t(n) = \eta p\sum_{m=1}^{n-1}\mu_t(m)\mu_t(n-m) - \eta\left(pq_t + (1-p)(1-q_t)\right)\mu_t(n), \quad \mu_0(n) = \delta_{n=1},\ \forall n\in A. \tag{260}$$

The equation in (260) has a solution of the form

$$\mu_t(n) = a_t^{n-1}b_t^{-1}, \quad n \in A, \tag{261}$$

which substituted in (260) yields

$$(n-1)a_t^{-1}\frac{\mathrm{d}}{\mathrm{d}t}a_t - b_t^{-1}\frac{\mathrm{d}}{\mathrm{d}t}b_t = -\eta\left(pq_t + (1-p)(1-q_t)\right) + \eta p(n-1)b_t^{-1}a_t^{-1}. \tag{262}$$

Separating variables and using (261) to write $q_t = \sum_{n\in A}a_t^{n-1}b_t^{-1} = b_t^{-1}\frac{a_t^{N-1}-1}{a_t-1}$, the Boltzmann equation in (260) decouples into the system of ODEs

$$\frac{\mathrm{d}}{\mathrm{d}t}b_t = \eta\left((1-p)b_t + (2p-1)\frac{a_t^{N-1}-1}{a_t-1}\right), \quad b_0 = 1, \tag{263}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}a_t = \eta p b_t^{-1}, \quad a_0 = 0. \tag{264}$$

To solve the system in (263), conjecture that $a_t \equiv G^{-1}(\eta pt + G(0))$ for some function $G(\cdot)$ to be determined. Substituting this conjecture in the last ODE in (263) immediately yields $b_t = G'(a_t)$.

Substituting this expression in the first ODE in (263) yields an ODE for $G'(x)$:

$$\mathrm{d}\log(G'(x)) = \frac{1-p}{p}G'(x) + \frac{2p-1}{p}\frac{x^{N-1}-1}{x-1}, \quad G'(0)=1, \;\; G(1)=0. \tag{265}$$

This equation has an explicit solution of the form

$$G(x) = \frac{p}{p-1}\log\left(1 - \frac{\frac{1-p}{p}\int_1^x \exp\left(\frac{2p-1}{p}\sum_{n\in A}\frac{z^n}{n}\right)\mathrm{d}z}{1+\frac{1-p}{p}\int_1^0 \exp\left(\frac{2p-1}{p}\sum_{n\in A}\frac{z^n}{n}\right)\mathrm{d}z}\right) \tag{266}$$

$$\equiv \frac{p}{p-1}\log\left(1 - \frac{\frac{1-p}{p}F(x)}{1+\frac{1-p}{p}F(0)}\right), \quad \forall x \in \mathbb{R}_+ \tag{267}$$

where the function $F(\cdot)$ is defined as

$$F(x) = \int_1^x \exp\left(\frac{2p-1}{p}\sum_{n\in A}\frac{1}{n}z^n\right)\mathrm{d}z, \quad \forall x \in \mathbb{R}_+. \tag{268}$$

The inverse of the function $G(\cdot)$ is thus given by

$$G^{-1}(y) = F^{-1}\left(\frac{\left(1-\exp\left(\frac{p-1}{p}y\right)\right)(p+(1-p)F(0))}{1-p}\right), \quad \forall y \in \mathbb{R}_+. \tag{269}$$

Substituting back into $a_t = G^{-1}(\eta pt + G(0))$ and differentiating (266) to get $b_t = G'(a_t)$ yields:

$$a_t = F^{-1}\left(\frac{p(1-\exp((p-1)\eta t)) + (1-p)F(0)}{1-p}\right) \quad \text{and} \quad b_t = \frac{pF'(a_t)}{p+(1-p)(F(0)-F(a_t))}. \tag{270}$$

Although Eq. (257) does not have an explicit solution for all $n \in B$, it can be solved through Fast Fourier Transform (FFT). Define the function

$$\varphi_t^A(\omega) = \sum_{n\in A} e^{\mathrm{i}\omega n}\mu_t(n) = b_t^{-1}\frac{e^{\mathrm{i}\omega} - e^{\mathrm{i}\omega N}a_t^{N-1}}{1-e^{\mathrm{i}\omega}a_t}, \quad \forall \omega \in \mathbb{R} \tag{271}$$

where $\mathrm{i} = \sqrt{-1}$. Further denote the Fourier transform of μ by $\varphi_t(\omega) = \int_{\mathbb{R}} e^{\mathrm{i}\omega n}\mathrm{d}\mu_t(n)$. Using (248) with $f(n) \equiv e^{\mathrm{i}\omega n}$ and (271), the Fourier transform φ must satisfy the Riccati equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\varphi_t = \eta\left(p(\varphi_t - 2\varphi_t^A)(2q_t + \varphi_t - 2\varphi_t^A - 1) + 2\varphi_t^A(q_t+\varphi_t) - q_t\varphi_t - 2(\varphi_t^A)^2 - \varphi_t^A\right), \quad \varphi_0 = e^{\mathrm{i}\omega}. \tag{272}$$

Using (270) and (271), I now define

$$\phi_t^A = \sum_{n\in A} n\mu_t(n) = \frac{1}{\mathrm{i}}\frac{\partial}{\partial\omega}\varphi_t^A(\omega)\bigg|_{\omega=0} = b_t^{-1}\frac{1+a_t^{N-1}(a_t(N-1)-N)}{(1-a_t)^2}. \tag{273}$$

Substituting $f(n) \equiv n$ and (273) in (248), it follows that the cross-sectional average number of signal $\phi_t = \sum_{n\in\mathbb{N}} n\mu_t(n)$ satisfies the ODE

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi_t = \eta\left((pq_t + (1-p)(1-q_t))\sum_{n\in A} n\mu_t(n) + ((1-p)q_t + p(1-q_t))\sum_{n\in B} n\mu_t(n)\right) \tag{274}$$

$$= \eta\left((2p-1)(2q_t-1)\phi_t^A + (p+q_t-2pq_t)\phi_t\right) \quad \phi_0 = 1. \tag{275}$$

The solution to this equation is given explicitly by

$$\phi_t = b_t^{-1}e^{\eta t}\left(1+\eta(2p-1)\int_0^t e^{-\eta s}b_s(2q_s-1)\phi_s^A \mathrm{d}s\right) \le \exp(\eta t). \tag{276}$$

Substituting (273) into this equation, I obtain (45).

F.3. Average trajectories of number of ideas

Finally, for the analysis it is useful to obtain the average trajectory of a manager i's number of ideas conditional on manager i holding $n_T^i = k$ ideas at the horizon date, i.e., $\mathbb{E}\left[n_t^i \middle| n_T^i = k\right]$. Applying Bayes' rule, first observe that

$$\mathbb{P}\left[n_t^i = m \middle| n_T^i = k\right] = \frac{\mathbb{P}[n_t^i = m]\mathbb{P}[n_T^i = k | n_t^i = m]}{\mathbb{P}[n_T^i = k]} = \frac{\mu_t(m)\rho_{T-t}(k;m)}{\mu_T(k)} \tag{277}$$

where ρ is the probability that manager i gets $k-m$ ideas by the horizon date conditional on holding m ideas at time t. To compute this probability, apply the result of Lemma F.1 to $g_s \equiv \sum_{n\in\mathbb{N}} f(n)\rho_{T-s}(n;m)$ for $s > t$ with initial condition, $g_t = m$, which yields:

$$\rho_{T-t}(k;m) = \mathbf{1}_{\{k\ge m\}}e^{t-(k-m+1)T}\left(\left(e^{\eta T}-1\right)^{k-m-1}\left(e^{\eta(T-t)}-1\right)\right)^{\mathbf{1}_{\{k-m\ge 1\}}} \tag{278}$$

under idea sharing, whereas under idea origination:

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho_s(k;m) = -\eta\rho_s(k;m) + \eta\rho_s(k;m-1), \tag{279}$$

with initial condition, $\rho_t(n;m) = \delta_{n=m}$, where $\delta_{n=m}$ is a Dirac mass at $n = m$. It then follows that the average trajectory of manager i's number of ideas is given by

$$\mathbb{E}\left[n_t^i \middle| n_T^i = k\right] = \frac{1}{\mu_T(k)}\sum_{m=1}^{k} m\mu_t(m)\rho_{T-t}(k;m), \tag{280}$$

which has a closed-form solution under idea sharing:

$$\mathbb{E}\left[n_t^i \middle| n_T^i = k\right] = \frac{1-e^{\eta T} + e^{k\eta(T-t)}\left(e^{\eta t}-1\right)^k\left(e^{\eta T}-1\right)^{1-k}}{1-e^{\eta(T-t)}}. \tag{281}$$

Appendix G. Proof of Proposition 8 (Density of $t-$statistics)

In this appendix, I derive the probability density function of conditional alpha $t-$statistics explicitly.

Consider first the probability that a manager i's conditional $t-$statistics in Eq. (38) is below some threshold $x \in \mathbb{R}$ conditional on holding n ideas at time t:

$$\mathbb{P}\left[t^i_{\alpha,t} \leq x \middle| n^i_t = n\right] = \mathbb{P}\left[\text{sign}(\widehat{\theta}^i_t)\frac{\sigma_S^2|k_t|}{\sigma_S^2 + o^c_t\phi_t}X_t \leq x \middle| n^i_t = n\right] \tag{282}$$

$$= \mathbb{P}\left[\widehat{\theta}^i_t < 0, X_t \geq -x\frac{\sigma_S^2 + o^c_t\phi_t}{\sigma_S^2|k_t|} \middle| n^i_t = n\right] + \mathbb{P}\left[\widehat{\theta}^i_t > 0, X_t \leq x\frac{\sigma_S^2 + o^c_t\phi_t}{\sigma_S^2|k_t|} \middle| n^i_t = n\right]. \tag{283}$$

Further conditioning on $n > \phi_t$ and using Eq. (240), we can write this expression as:

$$\mathbb{P}\left[t^i_{\alpha,t} \leq x \middle| n^i_t = n \geq \phi_t\right] = \mathbb{P}\left[X_t < -\frac{b(n)}{a_t(n)}\epsilon^i_t, X_t \geq -x\frac{\sigma_S^2 + o^c_t\phi_t}{\sigma_S^2|k_t|} \middle| n \geq \phi_t\right] \tag{284}$$

$$+ \mathbb{P}\left[X_t > -\frac{b(n)}{a_t(n)}\epsilon^i_t, X_t \leq x\frac{\sigma_S^2 + o^c_t\phi_t}{\sigma_S^2|k_t|} \middle| n \geq \phi_t\right] \tag{285}$$

$$= 2\mathbb{P}\left[\epsilon^i_t \leq \frac{a_t(n)}{b(n)}\frac{\sigma_S^2 + o^c_t\phi_t}{\sigma_S^2|k_t|}x, -x\frac{\sigma_S^2 + o^c_t\phi_t}{\sigma_S^2|k_t|} \leq X_t < -\frac{b(n)}{a_t(n)}\epsilon^i_t \middle| n \geq \phi_t\right] \tag{286}$$

$$= 2\int_{-\infty}^{\frac{a_t(n)}{b(n)}\frac{\sigma_S^2 + o^c_t\phi_t}{\sigma_S^2|k_t|}x}\left(\Phi\left(-\frac{b(n)}{a_t(n)}\frac{\epsilon}{\sqrt{\mathbb{E}[X_t^2]}}\right) - \Phi\left(-\frac{x}{\sqrt{\mathbb{E}[X_t^2]}}\frac{\sigma_S^2 + o^c_t\phi_t}{\sigma_S^2|k_t|}\right)\right)\phi(\epsilon)\mathrm{d}\epsilon \tag{287}$$

where the second equality follows from the symmetry of the normal distribution and from reorganizing. Repeating the same steps, but conditioning on $n < \phi_t$ similarly yields

$$\mathbb{P}\left[t^i_{\alpha,t} \leq x \middle| n^i_t = n < \phi_t\right] = \mathbb{P}\left[X_t \geq \max\left\{-x\frac{\sigma_S^2 + o^c_t\phi_t}{\sigma_S^2|k_t|}, -\frac{b(n)}{a_t(n)}\epsilon^i_t\right\} \middle| n < \phi_t\right] \tag{288}$$

$$+ \mathbb{P}\left[X_t \leq \min\left\{x\frac{\sigma_S^2 + o^c_t\phi_t}{\sigma_S^2|k_t|}, -\frac{b(n)}{a_t(n)}\epsilon^i_t\right\} \middle| n < \phi_t\right] \tag{289}$$

$$= 2\mathbb{P}\left[X_t \leq \min\left\{x\frac{\sigma_S^2 + o^c_t\phi_t}{\sigma_S^2|k_t|}, -\frac{b(n)}{a_t(n)}\epsilon^i_t\right\} \middle| n < \phi_t\right] \tag{290}$$

$$= 2\left(\begin{array}{c} \mathbb{P}\left[\epsilon^i_t > -\frac{a_t(n)}{b(n)}\frac{\sigma_S^2 + o^c_t\phi_t}{\sigma_S^2|k_t|}x \middle| n < \phi_t\right]\mathbb{P}\left[X_t \leq x\frac{\sigma_S^2 + o^c_t\phi_t}{\sigma_S^2|k_t|}\right] \\ +\mathbb{P}\left[\epsilon^i_t < -\frac{a_t(n)}{b(n)}\frac{\sigma_S^2 + o^c_t\phi_t}{\sigma_S^2|k_t|}x, X_t \leq -\frac{b(n)}{a_t(n)}\epsilon^i_t \middle| n < \phi_t\right] \end{array}\right) \tag{291}$$

$$= 2\left(\begin{array}{c} \left(1 - \Phi\left(-\frac{a_t(n)}{b(n)}\frac{\sigma_S^2 + o^c_t\phi_t}{\sigma_S^2|k_t|}x\right)\right)\Phi\left(\frac{x}{\sqrt{\mathbb{E}[X_t^2]}}\frac{\sigma_S^2 + o^c_t\phi_t}{\sigma_S^2|k_t|}\right) \\ + \int_{-\infty}^{-\frac{a_t(n)}{b(n)}\frac{\sigma_S^2 + o^c_t\phi_t}{\sigma_S^2|k_t|}x}\Phi\left(-\frac{b(n)}{a_t(n)}\frac{\epsilon}{\sqrt{\mathbb{E}[X_t^2]}}\right)\phi(\epsilon)\mathrm{d}\epsilon \end{array}\right). \tag{292}$$

Now observe that the density of manager i's conditional $t-$statistics conditional on holding n ideas at time t is obtained by differentiating Eq. (282) with respect to x:

$$\mathbb{P}[t^i_{\alpha,t} \in (x, x + \mathrm{d}x)|n^i_t = n] = \mathbf{1}_{n<\phi_t}\frac{\mathrm{d}}{\mathrm{d}x}\mathbb{P}\left[t^i_{\alpha,t} \leq x \middle| n^i_t = n < \phi_t\right] + \mathbf{1}_{n\geq\phi_t}\frac{\mathrm{d}}{\mathrm{d}x}\mathbb{P}\left[t^i_{\alpha,t} \leq x \middle| n^i_t = n \geq \phi_t\right]. \tag{293}$$

Differentiating the expressions in Eqs. (284) and (288) and reorganizing yields

$$\frac{\mathrm{d}}{\mathrm{d}x}\mathbb{P}\left[t^i_{\alpha,t} \le x \middle| n^i_t = n < \phi_t\right] = \frac{\mathrm{d}}{\mathrm{d}x}\mathbb{P}\left[t^i_{\alpha,t} \le x \middle| n^i_t = n \ge \phi_t\right] = l_t(x)\left(1 + \mathrm{erf}\left(\frac{a_t(n)}{\sqrt{2}b(n)}\frac{\sigma_S^2 + o^c_t\phi_t}{\sigma_S^2|k_t|}x\right)\right), \tag{294}$$

where $l_t(\cdot)$ denotes the Gaussian density defined in Eq. (40). It follows that the density of conditional $t-$statistics at time t satisfies

$$\mathbb{P}[t^i_{\alpha,t} \in (x, x + \mathrm{d}x)] = \sum_{n\in\mathbb{N}^*}\mu_t(n)\mathbb{P}[t^i_{\alpha,t} \in (x, x + \mathrm{d}x)|n^i_t = n] \tag{295}$$

$$= l_t(x)\left(1 + \sum_{n\in\mathbb{N}^*}\mu_t(n)\mathrm{erf}\left(\frac{a_t(n)}{\sqrt{2}b(n)}\frac{\sigma_S^2 + o^c_t\phi_t}{\sigma_S^2|k_t|}x\right)\right). \tag{296}$$

I finally use the Taylor series of the error function to write this density as

$$\mathbb{P}[t^i_{\alpha,t} \in (x, x + \mathrm{d}x)] = l_t(x)\left(1 + \sum_{n\in\mathbb{N}^*}\mu_t(n)\frac{2}{\sqrt{\pi}}\sum_{k\in\mathbb{N}}\frac{(-1)^k}{k!(2k+1)}\left(\frac{a_t(n)}{\sqrt{2}b(n)}\frac{\sigma_S^2 + o^c_t\phi_t}{\sigma_S^2|k_t|}x\right)^{2k+1}\right) \tag{297}$$

$$= l_t(x)\left(1 + \sum_{k\in\mathbb{N}}(-1)^k c(k)x^{2k+1}\sum_{n\in\mathbb{N}^*}\mu_t(n)R_t(n)^{2k+1}\right) \tag{298}$$

where the coefficients $c(k)$ are all positive, decreasing and satisfy:

$$c(k) = \tfrac{1}{\sqrt{\pi}}\tfrac{1}{k!(2k+1)}\left(\tfrac{1}{\sqrt{2}}\right)^{2k-1}, \quad k \in \mathbb{N}. \tag{299}$$

Appendix H. Proof of Proposition 9 (Distribution shift)

In this appendix I show that the cross-sectional distribution of $t-$statistics is shifted to the left when the distribution of number of ideas is symmetric or satisfies Corollary 1.

I start by computing the probability that a manager i's $t-$statistic is negative:

$$\mathbb{P}[t^i_{\alpha,t} \le 0] = \int_{-\infty}^{0} l_t(x)\left(1 + \sum_{n\in\mathbb{N}^*}\mu_t(n)\mathrm{erf}\left(\frac{x}{\sqrt{2}}R_t(n)\right)\right)\mathrm{d}x \tag{300}$$

$$= \frac{1}{2} - \frac{1}{\pi}\sum_{n\in\mathbb{N}^*}\mu_t(n)\tan^{-1}\left(\frac{n-\phi_t}{\sqrt{n}}\frac{\sqrt{\mathbb{E}[SR_t^2]}}{\sigma_S|k_t|}\right). \tag{301}$$

Proving that Eq. (43) holds is thus equivalent to proving:

$$\sum_{n\in\mathbb{N}^*}\mu_t(n)\tan^{-1}\left(\frac{n-\phi_t}{\sqrt{n}}\varphi_t\right) \equiv \Upsilon_t < 0, \tag{302}$$

where $\varphi_t > 0$ is positive at all finite times.

Assuming the functional form for μ_t in Corollary 1, I start by bounding $\tan^{-1}(\cdot)$ in Eq. (302)

by a piecewise linear function:

$$\tan^{-1}\left(\varphi_t \tfrac{n-\phi_t}{\sqrt{n}}\right) \leq \mathbf{1}_{n\leq\lfloor\phi_t\rfloor} \tfrac{\phi_t-n}{\phi_t-1} \tan^{-1}(\varphi_t(1-\phi_t)) + \mathbf{1}_{n>\lfloor\phi_t\rfloor} \min\left\{\tfrac{n-\phi_t}{\sqrt{\phi_t}}, \tfrac{\pi}{2}\right\} \equiv g(n), \quad n \in \mathbb{N}^* \ , \tag{303}$$

where the first term exploits that the support is bounded at 1 and the second term exploits both that the function is concave for all $n \geq \phi_t$ and thus bounded above by its first-order Taylor expansion and that $\tan^{-1}(\cdot)$ is bounded above at $\pi/2$. Although this bound can be tightened, it is sufficient to prove Eq. (302) and simple enough to compute explicit expressions: rewriting μ_t in Eq. (27) as

$$\mu_t(n) = \phi_t^{-n}(\phi_t - 1)^{n-1}, \tag{304}$$

and taking expectations over the function $g(\cdot)$ in Eq. (303) yields the explicit bound:

$$\Upsilon_t < \left(\frac{\phi_t - 1}{\phi_t}\right)^{\lfloor\phi_t\rfloor} \frac{\varphi_t \lfloor\phi_t\rfloor}{\sqrt{\phi_t}} \left(\begin{array}{c} \left(1 + \frac{\pi\sqrt{\phi_t}}{2\varphi_t} - \left\lceil \phi_t + \frac{\pi\sqrt{\phi_t}}{2\varphi_t} \right\rceil\right) \left(\frac{\phi_t-1}{\phi_t}\right)^{\left\lceil \phi_t + \frac{\pi\sqrt{\phi_t}}{2\varphi_t}\right\rceil - (\lfloor\phi_t\rfloor+1)} \\ +\sqrt{\phi_t}\left(\frac{\tan^{-1}(\varphi_t(1-\phi_t))}{\varphi_t(\phi_t-1)} + \frac{1}{\sqrt{\phi_t}}\right) \end{array} \right). \tag{305}$$

By assumption, $\eta > 0$, which implies that $\phi_t > 1$ and thus the term inside the bracket is nonpositive, from which the inequality in Eq. (302) follows.

Note that the result also obtains when μ_t is strictly symmetric at all times. Since the function $\tan^{-1}(\cdot)$ in Eq. (302) is odd and negative for all $n \leq \lfloor\phi_t\rfloor$, I define a new function $f : [\lfloor\phi_t\rfloor + 1, 2\lfloor\phi_t\rfloor] \to \mathbb{R}_+$ by reflecting $\tan^{-1}(\cdot)$ in Eq. (302) first horizontally about zero and then vertically about ϕ_t to obtain:

$$f(n) = \tan^{-1}\left(\varphi_t \tfrac{n-(2\lfloor\phi_t\rfloor+1-\phi_t)}{\sqrt{2\lfloor\phi_t\rfloor+1-n}}\right), \quad n = \lfloor\phi_t\rfloor + 1, \ldots, 2\lfloor\phi_t\rfloor. \tag{306}$$

The concavity of the skill-to-luck ratio then implies that

$$f(n) > \tan^{-1}\left(\varphi_t \tfrac{n-\phi_t}{\sqrt{n}}\right), \quad n = \lfloor\phi_t\rfloor + 1, \ldots, 2\lfloor\phi_t\rfloor. \tag{307}$$

Using this result I obtain a strict bound for Eq. (302):

$$\Upsilon_t < \sum_{n\in[1,\lfloor\phi_t\rfloor]\cup(2\lfloor\phi_t\rfloor,\ldots,\infty)} \mu_t(n)\tan^{-1}\left(\frac{n-\phi_t}{\sqrt{n}}\varphi_t\right) + \sum_{n\in(\lfloor\phi_t\rfloor,2\lfloor\phi_t\rfloor]} \mu_t(n) f(n), \tag{308}$$

at all finite times. Symmetry further implies that

$$\sum_{n\in[1,\lfloor\phi_t\rfloor]} \mu_t(n)\tan^{-1}\left(\frac{n-\phi_t}{\sqrt{n}}\varphi_t\right) + \sum_{n\in(\lfloor\phi_t\rfloor,2\lfloor\phi_t\rfloor]} \mu_t(n) f(n) = 0 \tag{309}$$

and that μ_t has zero mass beyond twice its mean:

$$\mu_t(n) = 0, \quad n \geq 2\lfloor\phi_t\rfloor, \tag{310}$$

since it is symmetric and its support is bounded from the left at 1. Hence, in the symmetric case Eq. (308) directly leads to Eq. (302).

Appendix I. Details on computations in Section 6

In this appendix I discuss computational details related to the equilibrium solution in the presence of fund flows and fees. For brevity, I simply pinpoint results that differ from the baseline model.

Note first that flows and fees do not affect the results of Proposition 1. However, they modify portfolio strategies. To see how, I go over the main steps of Appendix B in the presence of fees and flows. Maximizing manager i's expected utility over compensation in Eq. (48) implies:

$$\max_{\theta^i} \mathbb{E}\left[-\exp\left(-\gamma f((\tau+1)W_T^i - \tau B_T)\right)\middle|\mathscr{F}_t^i\right]. \tag{311}$$

The solution to this maximization problem is

$$\theta_{T-}^i \equiv \theta^i(\Psi, n, T-) = \frac{1}{f(\tau+1)\gamma}\left(o_{T-}(n)\right)^{-1}\lambda_{T-}^\top\Psi + \frac{\tau}{\tau+1}\left(\frac{\lambda_{1,T-}}{\lambda_{2,T-}}\Delta_{T-}^i + \widehat{\Theta}_{T-}^i\right). \tag{312}$$

which substituted in the value function yields the boundary condition

$$J(W, B, \Psi, n, T-) = -\exp\left(-\gamma f((\tau+1)W - \tau B) - \frac{1}{2}\left(o_{T-}(n)\right)^{-1}\Psi^\top\Lambda_{T-}\Psi\right). \tag{313}$$

The problem has an additional state variable, B, which applying Ito's lemma to Eq. (50) satisfies:

$$\mathrm{d}B_t = \widehat{\Theta}_t^c \mathrm{d}P_t = \begin{pmatrix} \frac{\lambda_{1,t}}{\lambda_{2,t}} & 1 \end{pmatrix}\Psi_t^i \mathrm{d}P_t \equiv \overline{\omega}_t^\top \Psi_t^i \mathrm{d}P_t. \tag{314}$$

The associated HJB equation now satisfies

$$0 = \max_{\theta^i}\left\{J_W A_Q \Psi^i\theta^i + \frac{1}{2}J_{WW}B_Q^2(\theta^i)^2 + B_Q B_\Psi(n^i)^\top J_{W\Psi}\theta^i + J_{WB}B_Q^2\overline{\omega}^\top\Psi\theta^i\right\} + J_t + J_\Psi^\top A_\Psi \Psi^i \tag{315}$$

$$+\frac{1}{2}\mathrm{tr}(J_{\Psi\Psi}B_\Psi(n^i)B_\Psi(n^i)^\top) + J_B(\Psi^i)^\top\overline{\omega}^\top A_Q\Psi^i + \frac{1}{2}J_{BB}B_Q^2(\Psi^i)^\top\overline{\omega}^\top\overline{\omega}\Psi^i + B_Q\overline{\omega}^\top\Psi^i B_\Psi(n^i)^\top J_{B\Psi} \tag{316}$$

$$+\eta(n^i)\mathbb{E}^{\mathscr{L}_t(\widehat{Y}^i,\Delta n^i)}\Big[J(W^i, B^i, \Psi^i + \sigma(n^i,\Delta n^i)\widehat{Y}^i, n^i + \Delta n^i, t) - J(W^i, B^i, \Psi^i, n^i, t)\Big]. \tag{317}$$

The first-order condition then yields the following portfolio policy:

$$\theta_t^i \equiv \theta_t(\Psi^i, n^i) = -\frac{J_W A_Q\Psi^i + B_Q B_\Psi(n^i)^\top J_{W\Psi} + J_{WB}B_Q^2\overline{\omega}^\top\Psi}{J_{WW}B_Q^2}. \tag{318}$$

Substituting back in the HJB equation, tedious derivations show the ansatz of Theorem B.1 becomes:

$$J(W, B, \Psi, n, t) = -\exp\left(-\gamma f((\tau+1)W - \tau B) - u_t(n) - \frac{1}{2}\left(\Psi^\top R_t(n) + R_t(n)^\top\Psi + \Psi^\top M_t(n)\Psi\right)\right), \tag{319}$$

where R and M satisfy the system of equations in Theorem B.1. The equation for u differs, but is irrelevant for portfolio strategies and is thus omitted. As a result, the solution for R and M is identical to Lemma B.3. Substituting these expressions in the optimal policy yields Eq. (51).

To obtain price coefficients I now go over the main steps in Appendix C. Aggregating first portfolios at the horizon date:

$$\int_0^1 \theta^i_{T-}\mathrm{d}i = \sum_{n\in\mathbb{N}} \mu_T(n)\frac{1}{\gamma f}\left(o_{T-}(n)\right)^{-1}\lambda_{T-}^\top \Gamma_{T-}(n)\Psi_T + \tau\overline{\omega}_{T-}\Psi_T = (\tau+1)\mathbf{1}^\star\Psi_T, \tag{320}$$

which yields the boundary conditions

$$\lambda_{1,T-} = \tfrac{o^c_{T-}\phi_T}{(\tau+1)(\phi_T o^c_{T-}+\sigma_S^2)} \quad\text{and}\quad \lambda_{2,T-} = -\gamma f \tfrac{o^c_{T-}\sigma_S^2}{(\phi_T o^c_{T-}+\sigma_S^2)}\ . \tag{321}$$

Similarly, aggregating portfolios at date t, Eq. (204) becomes:

$$\sum_{n\in\mathbb{N}} \mu_t(n)\left(A_{Q,t} - B_{Q,t}\left(o_t(n)\right)^{-1} B_{\Psi,t}(n)^\top \Lambda_t\right)\Gamma_t(n)\omega_t = -\frac{\lambda_{1,t}}{\lambda_{2,t}}\gamma f(\tau+1)B_{Q,t}^2, \tag{322}$$

since $\overline{\omega}_t^\top\omega_t = 0$, which yields, $\frac{\lambda_{1,t}}{\lambda_{2,t}} = -\frac{\phi_t}{\gamma f(\tau+1)\sigma_S^2}$, and thus

$$o^c_t = \left(\frac{1}{\sigma_\Pi^2} + \left(\frac{\phi_0}{\gamma f(\tau+1)\sigma_\Theta\sigma_S^2}\right)^2 + \left(\frac{1}{\sigma_\Theta\sigma_S^2\gamma f(\tau+1)}\right)^2\int_0^t\left(\frac{\mathrm{d}}{\mathrm{d}s}\phi_s\right)^2\mathrm{d}s\right)^{-1}. \tag{323}$$

Finally, spelling out the second equation of the system Eq. (211) becomes

$$\frac{\mathrm{d}}{\mathrm{d}t}\lambda_{2,t} = -k_t\lambda_{2,t}B_{Q,t} + B_{Q,t}^2\left(\lambda_{2,t}\frac{\phi_t o^c_t + \sigma_S^2}{o^c_t\sigma_S^2} + \gamma f\right), \tag{324}$$

the solution of which is Eq. (52).

Furthermore, going through the steps of Appendix D and simplifying yields

$$\widehat{\theta}^i_t = \frac{n^i_t - \phi_t}{f\gamma(\tau+1)(\sigma_S^2 + o^c_t\phi_t)}\left(\frac{\sigma_S^2(\tau+1)+\tau o^c_t\phi_t}{\sigma_S^2(\tau+1)}\Delta_t + f\gamma o^c_t\Theta_t\right) + \frac{\sqrt{n^i_t}}{f\gamma(\tau+1)\sigma_S}\epsilon^i_t \tag{325}$$

$$+\,\tau\frac{\sigma_S^2 k_t(f\gamma\sigma_S^2\sigma_\Theta(\tau+1)+k_t o^c_t\phi_t)}{(\sigma_S^2(\tau+1)(f\gamma\sigma_\Theta - k_t) - \tau k_t o^c_t\phi_t)^2}\left(\frac{\phi_t}{f\gamma\sigma_S^2(\tau+1)}\Delta_t - \Theta_t\right) \tag{326}$$

$$\equiv a_{\Delta,t}(n^i_t)\Delta_t + a_{\Theta,t}(n^i_t)\Theta_t + a_{\epsilon,t}(n^i_t)\epsilon^i_t, \tag{327}$$

where the second line corresponds to the function $H(\cdot)$ in Eq. (53). To compute informational alphas, I substitute the equilibrium solution in Eq. (86) to obtain

$$A_{Q,t} = \tau\begin{pmatrix}\frac{k_t o^c_t(f\gamma\sigma_S^4\sigma_\Theta(\tau+1)+k_t o^c_t\phi_t(2\sigma_S^2+o^c_t\phi_t))}{(\tau+1)(\sigma_S^2+o^c_t\phi_t)^2}\\ 0\end{pmatrix} + \begin{pmatrix}1\\ \gamma f o^c_t\end{pmatrix}\frac{\sigma_S^4 k_t o^c_t(k_t-(\tau+1)\sigma_\Theta\gamma f)}{(\sigma_S^2+o^c_t\phi_t)^2} \tag{328}$$

$$\equiv b_{\Delta,t}(n^i_t)\Delta_t + b_{\Theta,t}(n^i_t)\Theta_t, \tag{329}$$

and

$$B_{Q,t} = \frac{o^c_t|\sigma_S^2(\tau+1)(k_t - f\gamma\sigma_\Theta)+\tau k_t\phi_t o^c_t|}{(\tau+1)(\sigma_S^2+\phi_t o^c_t)}. \tag{330}$$

With these expressions can then compute $t-$statistics by simulations using Eq. (22).

References

Acemoglu, D., Bimpikis, K., Ozdaglar, A., 2010. Dynamics of information exchange in endogenous social networks. Working Paper .

Admati, A. R., Bhattacharya, S., Pfleiderer, P., Ross, S. A., 1986. On timing and selectivity. Journal of Finance 41, 715–730.

Admati, A. R., Ross, S. A., 1985. Measuring investment performance in a rational expectations equilibrium model. Journal of Business 58, 1–26.

Almazan, A., Brown, K. C., Carlson, M., Chapman, D. A., 2004. Why constrain your mutual fund manager? Journal of Financial Economics 73, 289–321.

Andrei, D., Cujean, J., 2017. Information percolation, momentum and reversal. Journal of Financial Economics 123, 441–670.

Back, K., 1992. Insider trading in continuous time. The Review of Financial Studies 5, 387–409.

Barras, L., Scaillet, O., Wermers, R., 2010. False discoveries in mutual fund performance: Measuring luck in estimated alphas. Journal of Finance 65, 179–216.

Basak, S., Pavlova, A., Shapiro, A., 2007. Optimal Asset Allocation and Risk Shifting in Money Management. Review of Financial Studies 20, 1583–1621.

Benabou, R., Laroque, G., 1992. Using privileged information to manipulate markets: Insiders, gurus, and credibility. The Quarterly Journal of Economics 107, 921–958.

Berk, J. B., Green, R. C., 2004. Mutual fund flows and performance in rational markets. Journal of Political Economy 112, 1269–1295.

Branikas, I., Hong, H., Xu, J., 2017. Buy side, sell side. Working Paper .

Brennan, M. J., Cao, H. H., 1997. International portfolio investment flows. Journal of Finance 52, 1851–80.

Burnside, C., Eichenbaum, M., Rebelo, S., 2010. Booms and busts: Understanding housing market dynamics. Working Paper .

Carhart, M. M., 1997. On persistence in mutual fund performance. Journal of Finance 52, 57–82.

Chen, J., Hong, H., Huang, M., Kubik, J. D., 2004. Does fund size erode mutual fund performance? the role of liquidity and organization. American Economic Review 94, 1276–1302.

Cheng, P., Scaillet, O., 2007. Linear-quadratic jump-diffusion modeling. Mathematical Finance 17, 575–598.

Chevalier, J., Ellison, G., 1999. Are some mutual fund managers better than others? cross-sectional patterns in behavior and performance. Journal of Finance 54, 875–899.

Christoffersen, S. E. K., Sarkissian, S., 2009. City size and fund performance. Journal of Financial Economics 92, 252–275.

Cohen, L., Frazzini, A., Malloy, C., 2008. The small world of investing: Board connections and mutual fund returns. Journal of Political Economy 116, 951–979.

Colla, P., Mele, A., 2010. Information linkages and correlated trading. Review of Financial Studies 23, 203–246.

Coval, J. D., Moskowitz, T. J., 2001. The geography of investment: Informed trading and asset prices. Journal of Political Economy 109, 811–841.

Crane, A., Koch, A., Michenaud, S., 2015. Institutional investor networks and governance mechanisms. Working Paper .

Cuoco, D., Kaniel, R., 2011. Equilibrium prices in the presence of delegated portfolio management. Journal of Financial Economics 101, 264–269.

DeMarzo, P. M., Vayanos, D., Zwiebel, J., 2003. Persuasion bias, social influence, and unidimensional opinions. Quarterly Journal of Economics 118, 909–968.

Detemple, J., Kihlstrom, R., 1987. Acquisition dinformation dans un modle intertemporel en temps continu. L'Actualit conomique 63, 118–137.

Duffie, D., Malamud, S., Manso, G., 2009. Information percolation with equilibrium search dynamics. Econometrica 77, 1513–1574.

Duffie, D., Malamud, S., Manso, G., 2015. Information percolation in segmented markets. Journal of Economic, forthcoming .

Duffie, D., Sun, Y., 2007. Existence of independent random matching. Annals of Applied Probability 17, 386–419.

Dybvig, P. H., Ross, S. A., 1985. Differential information and performance measurement using a security market line. Journal of Finance 40, 383–99.

Fama, E. F., French, K. R., 2010. Luck versus skill in the cross-section of mutual fund returns. Journal of Finance 65, 1915–1947.

Ferson, W. E., Chen, Y., 2015. How many good and bad funds are there, really? Working Paper .

Garcia, D., Strobl, G., 2011. Relative wealth concerns and complementarities in information acquisition. Review of Financial Studies 24, 169–207.

Glode, V., 2011. Why mutual funds underperform. Journal of Financial Economics 99, 546 – 559.

Goetzmann, W. N., Ingersoll, J., Spiegel, M., Welch, I., 2007. Portfolio performance manipulation and manipulation-proof performance measures. Review of Financial Studies 20, 1503–1546.

Gray, W., Crawford, S., Kern, A., 2012. Do fund managers identify and share profitable ideas? Working Paper .

Grinblatt, M., Titman, S., 1989. Mutual fund performance: An analysis of quarterly portfolio holdings. The Journal of Business 62, pp. 393–416.

Grossman, S. J., Stiglitz, J. E., 1980. On the impossibility of informationally efficient markets. American Economic Review 70, 393–408.

Gruber, M. J., 1996. Another puzzle: The growth in actively managed mutual funds. Journal of Finance 51, 783–810.

He, H., Wang, J., 1995. Differential informational and dynamic behavior of stock trading volume. Review of Financial Studies 8, 919–972.

Hong, D., Hong, H., Ungureanu, A., 2010. Word of mouth and gradual information diffusion in asset markets. Working Paper .

Hong, H., Kubik, J. D., Stein, J. C., 2005. Thy neighbor's portfolio: Word-of-mouth effects in the holdings and trades of money managers. Journal of Finance 60, 2801–2824.

Hugonnier, J., Kaniel, R., 2010. Mutual fund portfolio choice in the presence of dynamic flows. Mathematical Finance 2, 187–227.

Jazwinski, A., 1970. Stochastic Processes and Filtering Theory. Academic Press, New York.

Jensen, M. C., 1968. The performance of mutual funds in the period 1945-1964. Journal of Finance 23, 389–416.

Jones, R. C., Wermers, R., 2011. Active management in mostly efficient markets. Financial Analysts Journal .

Kaniel, R., Kondor, P., 2013. The delegated lucas tree. Review of Financial Studies 26, 929–984.

Karatzas, I., Shreve, S., 1988. Brownian Motion and Stochastic Calculus. Springer, Graduate Texte in Mathematics.

Koijen, R. S. J., 2012. The cross-section of managerial ability, incentives, and risk preferences. Journal of Finance .

Kosowski, R., Timmermann, A., Wermers, R., White, H., 2006. Can mutual fund "stars" really pick stocks? new evidence from a bootstrap analysis. Journal of Finance 61, 2551–2595.

Kothari, S., Warner, J., 2001. Evaluating mutual fund performance. Journal of Finance 56, 1985–2010.

Liptser, R. S., Shiryaev, A. N., 2000. Statistics of Random Processes I. Springer Verlag, New York.

Liptser, R. S., Shiryaev, A. N., 2001. Statistics of Random Processes II. Springer Verlag, New York.

Malkiel, B. G., 1995. Returns from investing in equity mutual funds 1971 to 1991. Journal of Finance 50, 549–572.

Mamaysky, H., Spiegel, M., Zhang, H., 2008. Estimating the dynamics of mutual fund alphas and betas. Review of Financial Studies 21, 233–264.

Ozsoylev, H., Walden, J., 2011. Asset pricing in large information networks. Journal of Economic Theory, forthcoming .

Pastor, L., Stambaugh, R. F., 2012. On the Size of the Active Management Industry. Journal of Political Economy 120, 740 – 781.

Pastor, L., Stambaugh, R. F., Taylor, L., 2016. Do funds make more when they trade more? Working Paper .

Pool, V. K., Stoffman, N., Yonker, S. E., 2015. The people in your neighborhood: Social interactions and mutual fund portfolios. The Journal of Finance 70, 2679–2732.

Schmidt, D., 2015. Stock market rumors and credibility. Working Paper .

Shiller, R. J., Pound, J., 1989. Survey evidence on diffusion of investment among institutional investors. Journal of Economic Behavior and Organization 12, 47–66.

Spiegel, M. I., Mamaysky, H., 2001. A theory of mutual funds: Optimal fund objectives and industry organization. Yale School of Management Working Papers ysm219, Yale School of Management.

Stein, J. C., 2008. Conversations among competitors. American Economic Review 98, 2150–62.

Treynor, J. L., Mazuy, K. K., 1966. Can mutual funds outguess the market ? Harvard Business Review 44, 131–136.

Vayanos, D., Woolley, P., 2011. Fund Flows and Asset Prices: A Baseline Model. FMG Discussion Papers dp667, Financial Markets Group.

Wu, Y., Wermers, R., Zechner, J., 2015. Managerial rents vs. shareholder value in delegated portfolio management: The case of closed-end funds. Working Paper .

Table 1
Summary of the notation. This table summarizes the main parameters (part 1), variables (part 2-4), and information sets (part 5) of the model.

Symbol	Definition
1. parameters of the model	
T	horizon date
γ	coefficient of absolute risk aversion
σ_Π	volatility of the fundamental
σ_Θ	volatility of the supply
σ_S	volatility of the noise in individual ideas
η	arrival rate of ideas
N	network threshold
τ	performance-flow parameter
f	fulcrum performance fee (rate)
2. equilibrium variables	
Π	fundamental value of the stock
Θ	(noisy) supply of the stock
$\widehat{X}^j$	expectation of X conditional on $\mathscr{F}^j$
o^j	posterior variance of the fundamental conditional on $\mathscr{F}^j$
Δ	informational advantage achieved under perfect information
P	equilibrium stock price
λ_1	price sensitivity to the fundamental
λ_2	price sensitivity to the supply
k	speed at which prices reveal information
3. idea-gathering variables	
n	number of ideas
μ	cross-sectional distribution of number of ideas
π	distribution of incremental number of ideas
ϕ	cross-sectional average number of ideas
4. performance variables	
W	wealth (assets under management)
θ	holdings in the stock
$\widehat{\theta}$	informational holdings (holdings net of per capital supply shocks)
α	informational alpha (alpha relative to average manager)
t_α	alpha $t-$statistic
t_{max}	maximal level of statistical significance
$s(n)$	skill (distance between n and the cross-sectional average number of ideas)
$R(n)$	skill-to-luck ratio (trading intensity on informatioin relative to noise)
SR	market Sharpe ratio
5. information sets	
$\mathscr{F}^c$	common information set (price history)
$\mathscr{F}^i$	information set of manager i (price history and collection of ideas)
$\mathscr{F}$	information set of the econometrician (fundamental, supply, informational holdings)

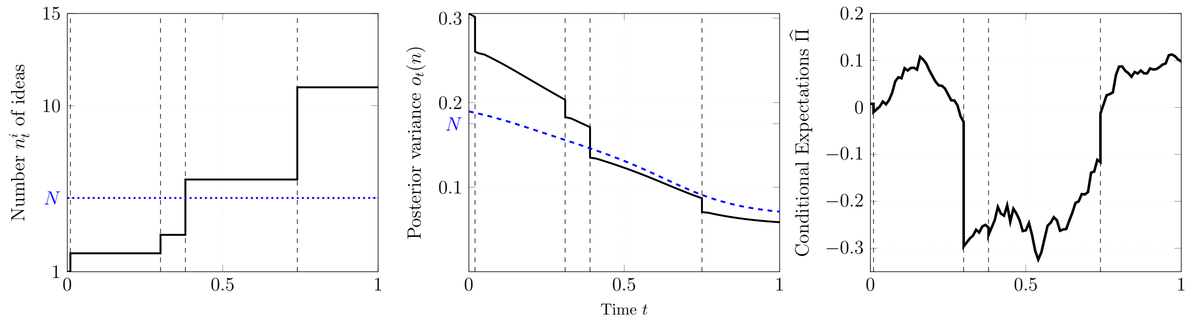

Fig. 1. Simulated Path of a Manager's Expectations. This figure plots a simulated path of a manager i's number n^i of ideas (left panel), her posterior variance o^i (the middle panel), and her expectations of the fundamental $\widehat{\Pi}^i$ (the right panel).

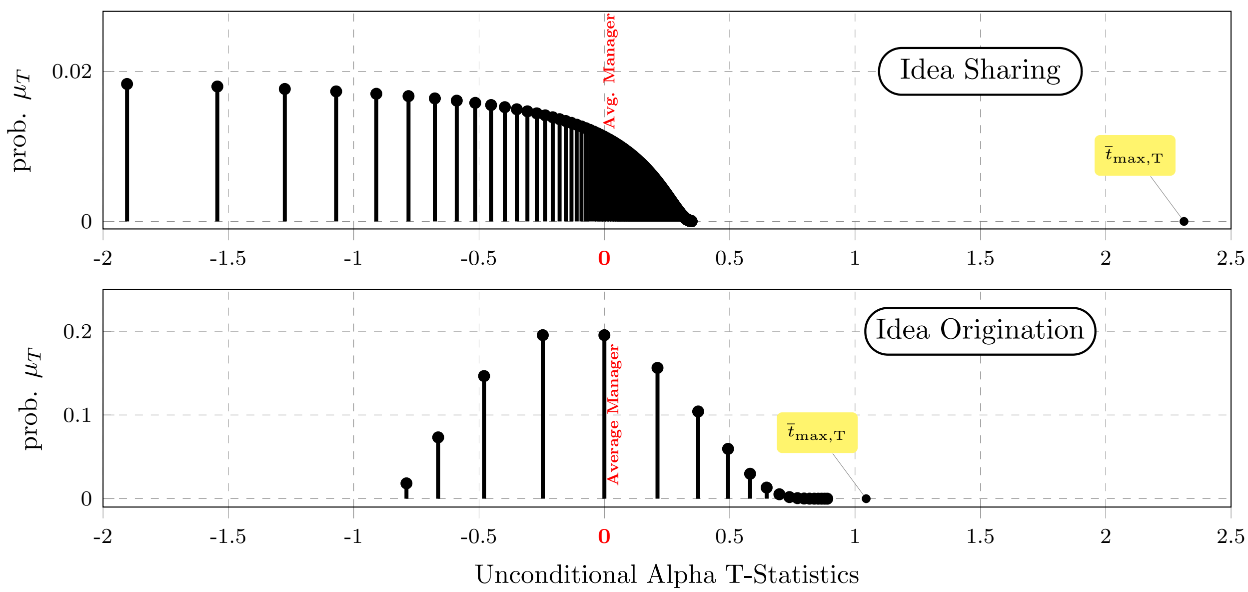

Fig. 2. Unconditional cross-sectional distribution of alpha t-statistics. The upper and the lower panel depict the cross-sectional distribution of unconditional alpha $t-$statistics under idea sharing and origination, respectively. The calibration is $\sigma_\Theta = \sigma_S = \sigma_\Pi = 1$, and $\gamma = 3$; it assumes a yearly sample size, $T = 1$, and $\eta = 4$ ideas a year on average. Each panel shows the average manager's performance and the maximal level of statistical significance, $\bar{t}_{\text{max,T}}$.

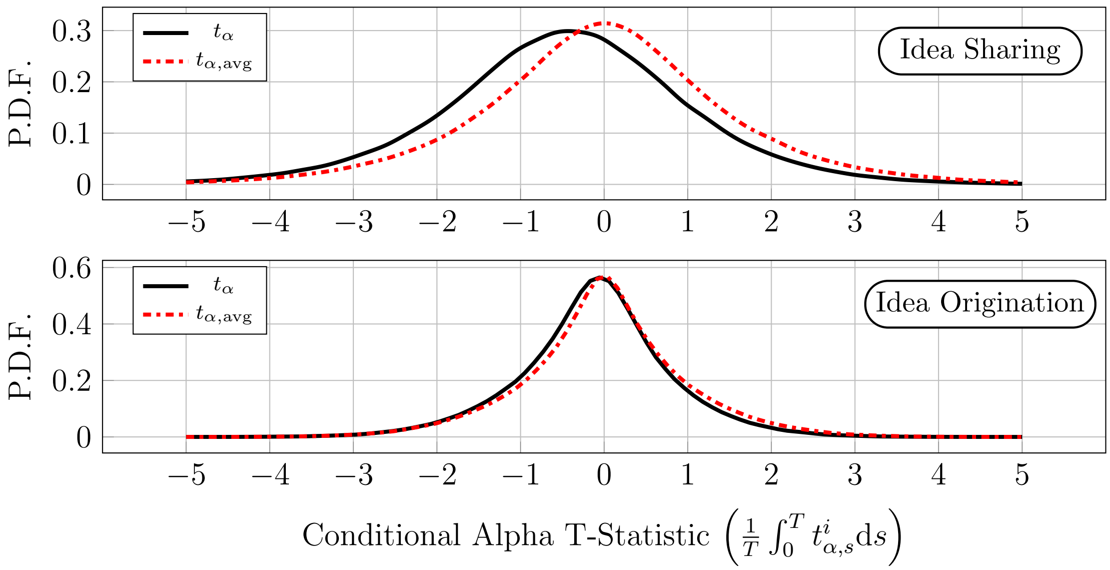

Fig. 3. Cross-sectional distribution of alpha t-statistics. The upper and lower panel plot the cross-sectional distribution of conditional alpha $t-$statistics averaged over the trading period (the solid black line) under idea sharing and origination, respectively. The calibration is $\sigma_\Theta = \sigma_S = \sigma_\Pi = 1$, and $\gamma = 3$; it assumes a yearly sample size, $T = 1$, and $\eta = 4$ ideas a year on average. Each panel shows the distribution generated by the average manager (the red dashed line). All distributions are obtained from 500000 simulations of the economy.

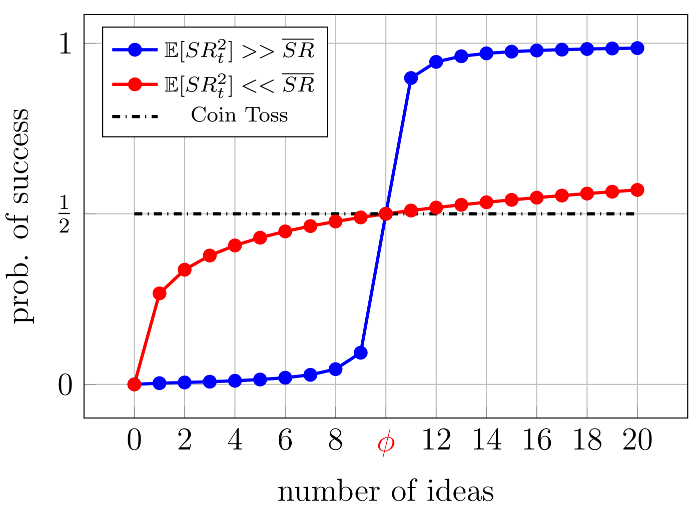

Fig. 4. Probability of successful market timing. This figure plots a manager's probability of successful market timing as a function of her number of ideas. The blue and the red line correspond to the case in which the squared Sharpe ratio is large and small, respectively. The dashed-dotted line represents pure luck.

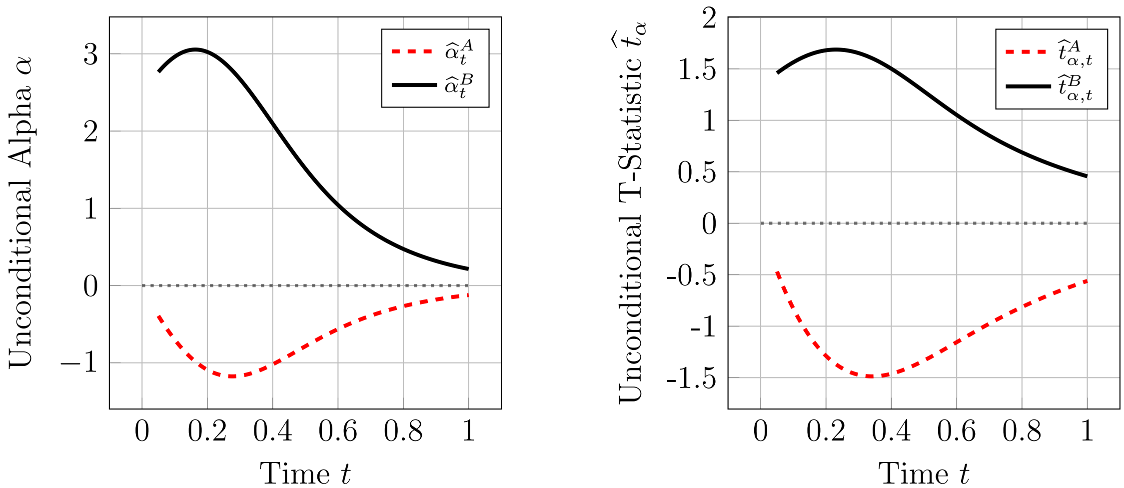

Fig. 5. Evolution of alpha and its t-statistic under exogenous segmentation. This figure plots the unconditional alpha (left panel) and its $t-$statistic (right panel) as a function of time when the population is exogenously split into unskilled (dashed red line, Group A) and skilled (solid black line, Group B) managers. The calibration is $\sigma_\Theta = \sigma_S = \sigma_\Pi = 1$, and $\gamma = 3$; it assumes a yearly sample size, $T = 1$, and $\eta = 4$ ideas a year on average.

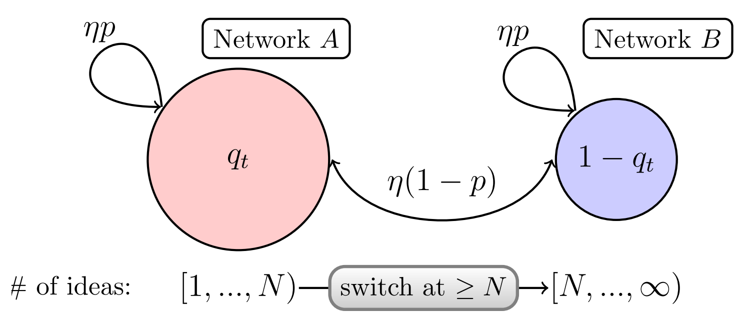

Fig. 6. Network Configuration. This figure illustrates the matching process and network formation.

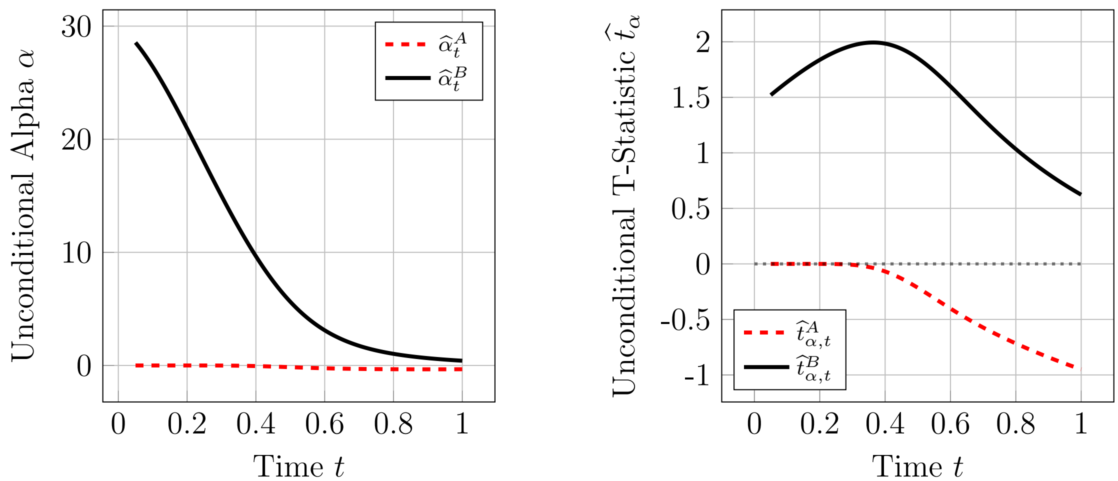

Fig. 7. Evolution of alpha and its t-statistic under network formation. This figure plots unconditional alpha (left panel) and its $t-$statistic (right panel) as a function of time when the population is endogenously segmented into Network A (dashed red line) and B (solid black line). The calibration is $\sigma_\Theta = \sigma_S = \sigma_\Pi = 1$, and $\gamma = 3$; it assumes a yearly sample size, $T = 1$, and $\eta = 4$ ideas a year on average, $N = 17$ and $p = 65\%$ of matches occuring in-network.

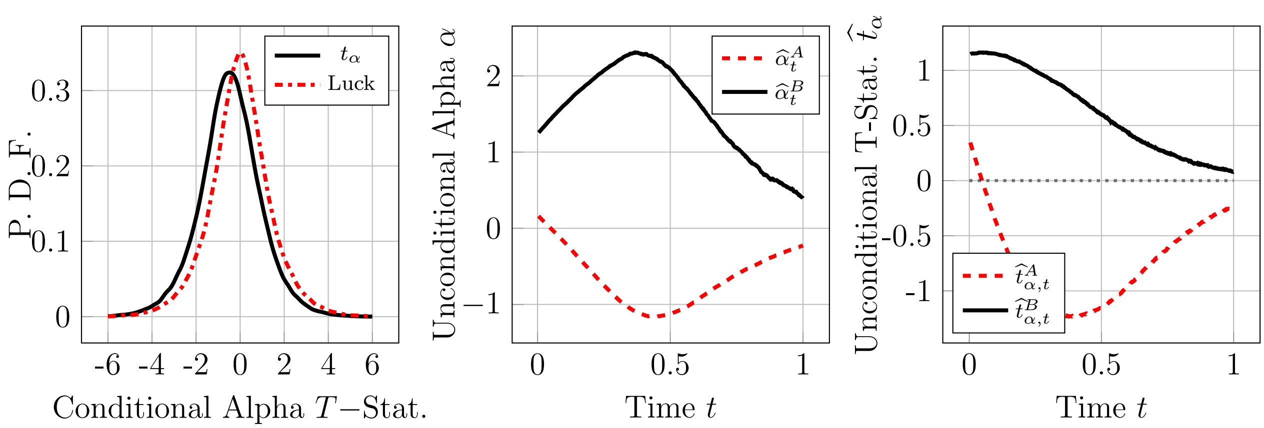

Fig. 8. Fund flows and performance under idea sharing. The left panel plots the distribution of average $t-$statistics under idea sharing with fund flows; the red dashed line is the distribution under pure luck. In the same setup, the middle and right panels plot the unconditional alpha and its $t-$statistic as a function of time. The dashed red line corresponds to the average unskilled manager (Group A) and the solid black line corresponds to the average skilled manager (Group B). The calibration is $\sigma_\Theta = \sigma_S = \sigma_\Pi = 1$, and $\gamma = 3$; it assumes a yearly sample size, $T = 1$, and $\eta = 4$ ideas a year on average, and $\tau = 0.86$, and all distributions are obtained from 200000 simulations.