NES New Economic School

Working Paper

No 287

NES Working Paper series

ADVERTISING ARBITRAGE

Sergei Kovbasyuk
Marco Pagano

February 2022

Advertising Arbitrage*

Sergey Kovbasyuk† Marco Pagano‡

February 2, 2022

Abstract

An arbitrageur with short investment horizon gains from accelerating price discovery by advertising his private information. However, advertising many assets may overload investors' attention, reducing the number of informed traders per asset and slowing price discovery. So the arbitrageur optimally concentrates advertising on just a few assets, unless his trades have significant price impact. The arbitrageur's gain from advertising is increasing in the assets' mispricing and in the precision of his private information, and is decreasing in its complexity. If several arbitrageurs have private information, inefficient equilibria can arise, where substantial mispricing persists or investors' attention is overloaded.

Keywords: limits to arbitrage, advertising, price discovery, limited attention.

JEL classification: G11, G14, G2, D84.

---

* We are grateful to an anonymous referee, Marco Bassetto, Bruno Biais, Elena Carletti, Eduardo Davila, Thierry Foucault, Mikhail Golosov, Antonio Guarino, Denis Gromb, Hugo Hopenhayn, Tullio Jappelli, Marcin Kacperczyk, Ralph Koijen, Peter Kondor, Guy Laroque, Alan Moreira, Marco Ottaviani, Nicola Pavoni, Nicola Persico, Andrea Pozzi, Wenlan Qian, Ioanid Rosu, Elia Sartori, Joel Shapiro, Jean Tirole and especially to Alexander Ljungqvist for insightful remarks and suggestions. We also thank participants in seminars at Bocconi, CSEF, EIEF, ICEF (HSE) Moscow, New Economic School (NES), Toulouse, UCL, Vienna, at the EFA and AEA annual meetings, the European Summer Symposium in Financial Markets, the 10th CSEF-IGIER Symposium on Economics and Institutions, and the Oxford Financial Intermediation Theory Conference for their comments. We acknowledge financial support from EIEF and NES.

†New Economic School, Moscow. E-mail: skovbasyuk@gmail.com.

‡University of Naples Federico II, CSEF and EIEF. E-mail: pagano56@gmail.com.

Introduction

Professional investors often "talk up their book." That is, they openly advertise their positions. Recently, some have taken not simply to disclosing their positions and expressing opinions, but to backing their assertions with data on the allegedly mispriced assets. Examples range from prominent hedge funds presenting their buy or sell recommendations on individual stocks at regular conferences attended by other institutional investors[1] to small investigative firms (like Muddy Waters Research, Glaucus Research Group, Citron Research and Gotham City Research) shorting companies while publishing evidence of fraudulent accounting and recommending "sell."[2]

This advertising activity is associated with abnormal returns: Ljungqvist and Qian (2016) examine the reports that 31 professional investors published upon shorting 124 US listed companies between 2006 and 2011, and find that they managed to earn substantial excess returns on their short positions, especially when the reports contained hard information. Luo (2018) finds that the stocks pitched by hedge funds at conferences – mostly with "buy" recommendations – perform better than other stocks held by the same funds, earning abnormal returns both in the 18 months before and in the 9 months after the pitch. In the context of social media, Chen et al. (2014) document that articles and commentaries disseminated by investors via the social network Seeking Alpha predict future stock returns, witnessing their influence on the choices of other investors and so eventually on stock prices.

These examples tell a common story: professional investors who detect mispriced securities ("arbitrageurs") often advertise their information in order to accelerate the correction. Without such advertising, prices might diverge even further from fundamentals, whereas successful advertising will push prices closer to fundamentals, and enable the arbitrageurs to close their positions profitably. To make sure their advertising is successful, these ar-

[1] The best known such events are the Robin Hood, the Sohn Investment and the SkyBridge Alternatives Conference. Robin Hood (www.robinhood.org) attendees typically pay 7,500 dollars or more to the charity for a ticket, although of course their attendance is largely motivated by the desire to be the first to hear the hedge fund managers' pitches: often, in fact, they trade on them from their smartphones while the conference is still in session. Famous examples of advertising campaigns run by large hedge funds include David Einhorn's Greenlight Capital talking down and shortselling the shares of Allied Capital, Lehman Brothers and Green Mountain Coffee Roasters.

[2] For instance, in July 2014 Gotham City Research provided evidence of accounting fraud in the Spanish company Gowex, causing its stock price to collapse and forcing the company to file for bankruptcy: see *The Economist*, "Got'em, Gotham", 12 July 2014, p.53.

bitrageurs typically go well beyond simply stating their recommendation: they produce hard evidence buttressing it during their pitches, and typically disclose their positions to impart additional credibility.[3] This mechanism is crucial for arbitrageurs with sizable holding costs per unit of time, such as short sellers, who need to finance margin requirements, but it is also relevant for those with long positions, insofar as they seek high short-term returns. That short-termism is a key determinant of such advertising activity is consistent with the evidence reported by Pasquariello and Wang (2021).

In this paper, we present a model in which risk-averse arbitrageurs can advertise their private information about mispriced assets to rational investors with limited attention, and at the same time choose their portfolios to exploit the price correction induced by such advertising. Under the convenient assumptions of constant-absolute risk aversion (CARA) and normal fundamentals and noise trading, we show that insofar as advertising succeeds in catching the attention of rational investors, it reduces the risk incurred by the arbitrageur in liquidating his position due to noise traders. This risk reduction in turn enhances the arbitrageur's willingness to make large bets on his private information, engendering a complementarity between advertising and investing in the advertised securities.[4] Owing to the interaction between these two choices, the model yields a number of predictions about advertising activity, arbitrageurs' portfolio choices and equilibrium prices, some consistent with the evidence provided by recent studies and others still to be tested.

First, even when an arbitrageur identifies a number of mispriced assets, he will concentrate his advertising on just a few:[5] diluting investors' attention across too many assets would reduce the number of informed traders for each, diminish price discovery and leave a large liquidation risk for all. That is, concentrated advertising is a safer bet than diversified advertising: it increases the chances of closing the position profitably. Indeed, in practice

[3]Hence such advertising differs from the release of soft information by market gurus, who cannot justify their trading recommendations with hard information. Benabou and Laroque (1992) show that gurus can manipulate the market by issuing negative messages about high-value assets, buying them cheap and reselling them at high prices later on. Yet, this requires gurus to have a long horizon, since, as shown by Schmidt (2019), short-horizon speculators will tell the truth even if their messages are cheap talk.

[4]In contrast, trading and communication with other investors are substitutes in the model by Liu (2017), where a large informed short-term investor can reveal his information either through trading or via communication, and under some circumstances will prefer to communicate and trade less aggressively, rather than remain silent and rely solely on trading.

[5]This parallels the result in Lipnowski et al. (2020) where a principal who has complex information and faces an agent with limited attention optimally engages in "attention management", in the sense that he "restricts some information to induce the agent to pay attention to other aspects."

hedge fund managers who advertise their recommendations typically pitch a single asset at a time.

Second, concentrated advertising produces portfolio under-diversification. By lowering liquidation risk, advertising a given mispriced asset raises the arbitrageur's risk-adjusted expected return, inducing him to overweight that asset in his portfolio.

Thirdly, if an arbitrageur has private information about a number of assets, he will get the most out of his advertising if he pitches those for which mispricing is largest, his private information is most precise, and there is least noise trading. Moreover, the arbitrageur will prefer to advertise assets whose information is easy to process, as the corresponding ads require less attention by investors.

Fourthly, multiple arbitrageurs with common information will exhibit "wolf pack" behavior, advertising and trading the same assets. Intuitively, no individual arbitrageur has the incentive to deviate and divert investors' attention to assets not advertised by others, because this would lower his returns from assets already advertised by others and lower total expected payoff. Hence, in equilibrium each arbitrageur mimics the others. However, this "piggybacking" also tends to generate multiple equilibria, some of which are inefficient. For instance, arbitrageurs may get collectively trapped in a situation where they all advertise assets that are not the most sharply mispriced. This may explain why at times the market appears to pick up minor mispricing of some assets and neglect much more pronounced mispricing of others, such as RMBSs and CDOs before the financial crisis. Our multiple equilibria echo those found by Froot et al. (1992), where short-horizon speculators choose to learn the same signal as others, even if it has little or no informational content. In our model the inefficiency arises not from arbitrageurs' learning decisions, but from strategic complementarity in their choices about which signals to advertise to investors, namely, their incentive to avoid diluting investors' limited attention with heterogeneous information.

Instead, if arbitrageurs have exclusive private information about different assets, they always overload investors' attention to the extreme, so that collectively they have lower expected payoff than when they have common information. As they cannot coordinate, they end up over-exploiting investors' attention, just as in the tragedy of the commons: advertising too many assets leads to too little attention being paid to each, hence too much persistence of mispricing.

Finally, we explore how the results change when arbitrageurs are large relative to the market, so that their trades have price impact, both initially when they build their position and subsequently when they liquidate it. In both cases, price impact reduces the profitability of arbitrage, and therefore affects the arbitrageur's optimal advertising strategy. In this case we find that the result that the arbitrageur refrains from over-exploiting investors' attention capacity is attenuated: now there is a parameter region where the arbitrageur will want to advertise more assets than an individual investor can process. Intuitively, this is because the desire to reduce price impact may push the arbitrageur to spread his trades across many assets, thus mitigating the price impact for each of them.

Our model spans two strands of research: the literature on limited attention in asset markets, which studies portfolio choice and asset pricing when investors cannot process all the relevant information (Barber and Odean (2008), DellaVigna and Pollet (2009), Huberman and Regev (2001), Peng and Xiong (2006), and Van Nieuwerburgh and Veldkamp (2009, 2010)), and that on the limits to arbitrage and its inability to eliminate all mispricing (see Shleifer and Vishny (1997) and Gromb and Vayanos (2010), among others). In our setting, advertising informs investors, thereby reducing mispricing, but investors' limited attention forces the arbitrageur to restrict advertising to a few assets. Advertising also adds a dimension that is lacking in the limits-to-arbitrage models: it enables arbitrageurs to effectively relax those limits and endogenously speed up the movement of capital towards arbitrage opportunities.

Two of our results are reminiscent of those produced by other models, although they stem from a different source. First, in our setting arbitrageurs, like investors in Van Nieuwerburgh and Veldkamp (2009, 2010) and Veldkamp (2011), choose under-diversified portfolios but for a different reason. Our arbitrageurs have unlimited information-processing capacity (and may be informed about several arbitrage opportunities), so that in principle they could choose well-diversified portfolios. Instead they choose under-diversified portfolios for efficiency in advertising: the limited attention of their target investors affects their own portfolio choices, and biases them towards the advertised assets. Our paper also differs from the setup in Van Nieuwerburgh and Veldkamp (2010) because we model limited attention capacity as a ceiling on the number of signals that an investor can process, rather than as a bound on the total precision of all processed signals. However, in an extension we show that our assumption that an investor can process a limited number of signals can be obtained in the entropy-based framework considered in Van Nieuwerburgh and Veldkamp (2010).

Second, the herd behavior that arises in the presence of multiple arbitrageurs is superficially reminiscent of what happens in models of informational cascades such as Scharfstein and Stein (1990) and Bikhchandani et al. (1992). But in our model herding arises from the strategic complementarity between advertising and investing by arbitrageurs, and speeds up price discovery. By contrast, in informational cascades investors disregard their own information in favor of inference from the behavior of others, which tends to delay price discovery.

Our analysis of the interactions among arbitrageurs can also be related to Abreu and Brunnermeier (2002), who argue that arbitrage may be delayed by synchronization risk: in their model, arbitrageurs learn about an opportunity sequentially, and thus prefer to wait when they are unsure that enough of them have learned of it to correct the mispricing. Abreu and Brunnermeier (2002) hypothesize that announcements – like advertising in our model – may facilitate coordination among arbitrageurs and accelerate price discovery. In our model, by contrast, when mispricing is known to a number of arbitrageurs, there is no synchronization risk, yet advertising may not help eliminate the most acute mispricing, because of multiple equilibria.

The paper is organized as follows. Section 1 lays out the model with a single informed arbitrageur. Section 2 derives investors' portfolio and information processing choices and the resulting equilibrium prices of assets, taking the decisions of the arbitrageur as given. Section 3 characterizes the arbitrageur's optimal advertising and investment decisions, and studies how asset characteristics affect the gain from advertising them. Next, Section 4 considers multiple informed arbitrageurs, allowing for strategic interactions among them. Finally, Section 5 relaxes an important assumption maintained up to that point, namely that arbitrageurs' trades have no price impact, and derives the conditions under which the result of concentrated advertising still holds. Section 6 summarizes and discusses our predictions.

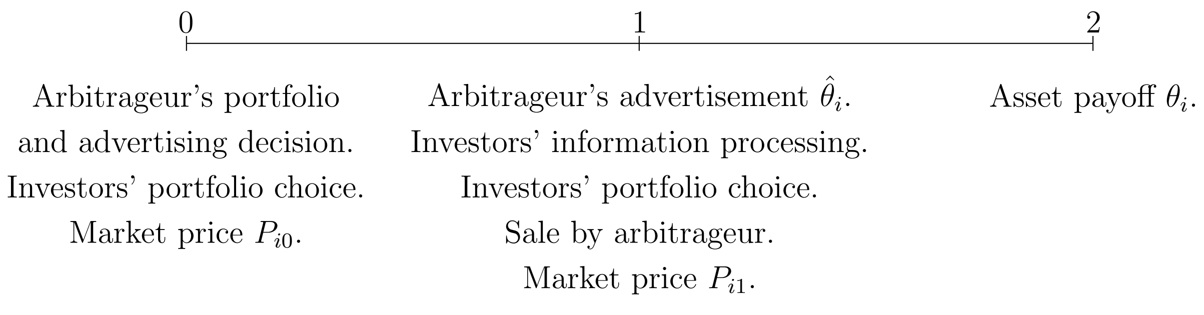
1 The Model

We consider an economy with a set $\mathbf{N}$ of risky assets, each available in zero net supply, and a safe asset that for simplicity is assumed to pay zero interest. The number of assets in this set is very large ($N \to \infty$). All assets are traded competitively by a unit mass of rational atomistic investors and by noise traders. Some of the rational investors have private information about a set of mispriced assets, which they can exploit. We refer to these investors as *arbitrageurs*. Initially, we consider the case of a single arbitrageur. In Section 4 we extend the analysis to multiple arbitrageurs, and in Section 5 we relax the assumption of price-taking behavior by arbitrageurs.

Timing. There are three periods: $t = 0, 1, 2$. As shown in Figure 1, each asset $i \in \mathbf{N}$ is traded at dates $t = 0, 1$ at prices P_{i0} and P_{i1} respectively, and at $t = 2$ delivers a final payoff θ_i with a normal distribution $N(\mu_i, \sigma^2_{\theta_i})$.

Investors. Different cohorts of investors trade in periods $t = 0$ and $t = 1$, and each cohort has a one-period horizon. Investors have constant absolute risk-aversion (CARA) preferences: their utility from a monetary payoff c is $1 - e^{-\rho c/2}$. The parameter $\rho/2 > 0$ is the Arrow-Pratt measure of absolute risk aversion. The total demand by noise traders for each asset $i \in \mathbf{N}$ is denoted by u_{i0} at $t = 0$ and u_{i1} at $t = 1$, where $u_{it} \sim N(0, \sigma^2_{u_i})$.

Figure 1: Timeline for each asset $i \in \mathbf{N}$



Information structure. At $t = 0$ the arbitrageur learns a private signal about the future payoff of a finite subset of M assets $i \in \mathbf{M} \subseteq \mathbf{N}$, which he can advertise to investors at $t = 1$. We denote this private signal by $\hat{\theta}_i = \theta_i + \varepsilon_i$, where $\varepsilon_i \sim N(0, \sigma^2_{\epsilon_i})$. The precision of the arbitrageur's signal is denoted by $\tau_{Ai} = 1/\sigma^2_{\epsilon_i}$, $i \in \mathbf{M}$. The arbitrageur has no private information about assets that do not belong to the set $\mathbf{M}$. Like other investors, the arbitrageur has a one-period horizon: if he takes a non-zero position x_i in asset i at $t = 0$, he must liquidate his position at $t = 1$.[6]

Other rational investors are unaware of where the arbitrageur's informational advantage lies: they do not know either the set $\mathbf{M}$ or the arbitrageur's signals, unless they learn them from advertising. From their point of view, any asset $i \in \mathbf{N}$ can be in $\mathbf{M}$ with the same probability, and since the number of traded assets is very large ($N \to \infty$), this probability is zero. Hence, unlike standard models of informed trading, this model posits no learning from prices: investors can learn about fundamentals only from arbitrageurs' advertising.

From the viewpoint of investors, the fundamental value of asset i is normally distributed with mean μ_i, i.e. $\theta_i = \mu_i + \eta_i$, and $\eta_i \sim N(0, \sigma^2_{\theta_i})$. Hence, the precision of investors' prior is $\tau_{\theta_i} \equiv 1/\sigma^2_{\theta_i}$.

Advertising. At $t = 0$ the arbitrageur may take positions in one or more of the M assets on which he has information, and then at the beginning of $t = 1$ he may advertise his private information about a subset of assets $\mathbf{L} \subseteq \mathbf{M}$ just before trading, in order to affect their valuations and thus their market prices. We denote by L the number of assets that he advertises.

As we shall see below, the arbitrageur's optimal position in an advertised asset x_i is proportional to his private signal $\hat{\theta}_i$. Hence, disclosing his position does not add any information to the signal. As explained above, investors do not learn from market prices either, because they do not know where the arbitrageur's informational advantage lies. Hence, investors who do not process the signal advertised by the arbitrageur must rely solely on their prior in their portfolio choices.

Limited attention. We assume that there is a limit to the amount of information an investor can process. This limit depends on the investor's total information capacity, denoted by C, and on the difficulty of processing arbitrageurs' ads, denoted by D, which can be thought as the fraction of complex information relative to simple information encoded in the ads. As a result, the effective information capacity of investors is $K = \lfloor C/D \rfloor$,

[6] One can think of the arbitrageur as incurring holding costs, reflecting the urgency of investing in other profitable assets, as in Abreu and Brunnermeier (2002).

which measures (up to the lowest integer) the maximum number of ads that an investor can process. If the arbitrageur advertises $L > K$ assets, each investor is assumed to receive and process a random sample of K ads, as before processing them investors are unaware of the characteristics of the asset corresponding to each ad, for instance its level of mispricing or return volatility. As we shall see, the model can also accommodate the case where ads concerning different assets feature different difficulty: denoting by D_i the difficulty of processing the information regarding asset $i \in \mathbf{L}$, the investor's information capacity must satisfy $C \geq \sum_{i \in \mathbf{K}} D_i$, where $\mathbf{K}$ is the set of assets for which the investor processes ads.

In our setting each individual investor is assumed not to split his attention capacity K across more than K assets, even when these are advertised. This is consistent with optimal allocation of limited attention capacity in Van Nieuwerburgh and Veldkamp (2010) when investors have mean-variance preferences, as in our setting (see their Proposition 2 and specifically Corollary 2): they find that in this case an investor with entropy learning technology does not want to split attention across independent assets. For illustrative purposes, in the Online Appendix we show that in our setting this result obtains when each investor has information processing capacity $K = 1$ and two assets are advertised ($L = 2$): he will prefer to learn perfectly about a single asset rather than imperfectly about two.

In what follows, we denote by m_i the mass of investors who learn from the ad of asset i as "informed" and to the remaining $1 - m_i$ investors as "uninformed". For instance, if the arbitrageur advertises $L > K$ assets featuring equal difficulty D, then each investor processes a random sample of K ads in $\mathbf{L}$, so that each ad is processed with probability $K/L < 1$. Hence, for each advertised asset the fraction of informed investors is $m_i = K/L < 1$.

Since advertising gives information for free to investors, one may wonder if the arbitrageur might not gain more by selling his information. But, as our analysis will show, the arbitrageur gains by disseminating his information to as many investors as possible: hence, he has no interest in limiting their number by charging for it. Moreover, information sales are difficult because to convince investors to buy his information the arbitrageur may have to disclose it, at which point investors would not be willing to pay for it – the well-known Arrow information paradox.

In what follows, we shall first derive the optimal portfolios of informed and uninformed investors and the equilibrium prices of assets, taking the arbitrageur's advertising decision as given. Next, we shall solve for the arbitrageur's optimal advertising and investment decisions.

2 Investors' Decisions

Consider asset $i \in \mathbf{M}$, i.e., one that can be advertised by the arbitrageur. In this section assets are assumed to be symmetric, so that the index i can be dropped with no loss of clarity. For instance, we shall refer to the mass of investors informed about asset i at $t = 1$ as $m \in [0, 1]$. In subsequent sections we shall extend the analysis to the heterogeneous assets case.

First, we focus on the determination of the prices of advertised assets at $t = 1$. Next, we turn to the determination of asset prices at $t = 0$.

2.1 Equilibrium prices at $t = 1$

The price P_1 must clear the market for the asset, balancing the net demand of the arbitrageur, noise traders, informed investors, and uninformed investors.

Investors' demand for the asset at $t = 1$ results from their optimal attention allocation decision as well as their portfolio choice. Attention allocation is trivial: if the arbitrageur advertises $L \leq K$ assets, investors can process signals about all the advertised assets, so that the mass of investors informed about each of them is $m = 1$. If instead the arbitrageur advertises $L > K$ assets, each investor randomly picks K assets and processes their respective ads, so that the mass of investors informed about each advertised asset is $m = K/L < 1$. Hence, the fraction of informed investors is

$$m(L) = \min[1, K/L]. \tag{1}$$

Taking the attention allocation decision and the resulting fraction $m \in (0, 1]$ of informed investors as given, we derive investors' portfolio choices and characterize the market clearing price for advertised assets. As an investor can be informed about some assets and uninformed about others, we denote his information set by Ω, which includes the advertised information processed by him. Since utility is CARA and asset payoffs are normal and independent, the investor's expected utility maximization is equivalent to a mean-variance optimization problem:

$$\max_{\{y_{i1}\}_{i\in N}} \sum_{i=1}^{N} \left[y_{i1}(\mathbb{E}(\theta_i|\Omega) - P_{i1}) - \frac{\rho}{2} y_{i1}^2 \mathbb{V}(\theta_i|\Omega) \right]. \tag{2}$$

If an asset is not advertised, all investors (except the arbitrageur) are uninformed about it. For advertised assets, instead, some investors will be informed and others uninformed. Informed investors condition their demand y_I on the arbitrageur's signal $\hat{\theta} = \theta + \varepsilon$, so that from their viewpoint the conditional distribution of the asset's future payoff at $t = 1$ is $N\left(\mathbb{E}(\theta|\hat{\theta}), \mathbb{V}(\theta|\hat{\theta})\right)$, where

$$\mathbb{E}(\theta|\hat{\theta}) = \frac{\mu\sigma_\varepsilon^2 + \sigma_\theta^2\hat{\theta}}{\sigma_\theta^2 + \sigma_\varepsilon^2} = \frac{\tau_\theta\mu + \tau_A\hat{\theta}}{\tau_\theta + \tau_A},$$

$$\mathbb{V}(\theta|\hat{\theta}) = \frac{\tau_\theta E[(\theta-\mu)^2] + \tau_A E[(\theta-\hat{\theta})^2]}{\tau_\theta + \tau_A} = \frac{1}{\tau_\theta + \tau_A}.$$

Each informed investor will maximize the certainty equivalent of the payoff from the asset:

$$\max_{\{y_I\}} \left(\frac{\tau_\theta\mu + \tau_A\hat{\theta}}{\tau_\theta + \tau_A} - P_1 \right) y_I - \frac{\rho}{2} \frac{y_I^2}{\tau_\theta + \tau_A}. \tag{3}$$

Hence, each of them buys

$$y_I = \frac{\tau_\theta\mu + \tau_A\hat{\theta} - (\tau_\theta + \tau_A)P_1}{\rho}. \tag{4}$$

If the asset is not advertised, investors form their expectations only based on their prior information, so that from their viewpoint the distribution of the asset's payoff is $N(\mu, 1/\tau_\theta)$. For uninformed investors, the probability that an asset is advertised at $t = 1$ is negligible, being equal to $L/N \to 0$, with $N \to \infty$. Therefore for them the distribution of the asset's future payoff coincides with the distribution of a non-advertised asset $N(\mu, 1/\tau_\theta)$. Hence, investors who are uninformed about an asset maximize the following certainty equivalent of their payoff from the asset:

$$\max_{\{y_U\}} \ (\mu - P_1)y_U - \frac{\rho}{2}\frac{y_U^2}{\tau_\theta}, \tag{5}$$

so that their demand for the asset is

$$y_U = \frac{\tau_\theta(\mu - P_1)}{\rho}. \tag{6}$$

The market will balance the demand from the m informed investors, the $1-m$ uninformed ones and the noise traders' demand u_1 (the arbitrageur's trade being negligible):

$$(1-m)y_U + my_I + u_1 = 0,$$

because the asset is in zero net supply. The resulting market clearing price is

$$P_1 = \frac{m\tau_A\hat{\theta} + \tau_\theta\mu}{m\tau_A + \tau_\theta} + \frac{\rho}{m\tau_A + \tau_\theta}u_1. \tag{7}$$

This expression indicates that in equilibrium the market learns from the arbitrageur's advertising, and that such learning is greater if the arbitrageur is regarded as being well informed (high τ_A), for instance because of a good track record: if his private information is thought to be precise, it gets a larger weight in price formation, and the price impact of noise trading is correspondingly reduced.

Being atomistic, the arbitrageur will be able to liquidate at price (7) whatever initial position he may have acquired at $t = 0$.[7] For assets that are not advertised, the equilibrium price at $t = 1$ is obtained by setting $m = 0$ in expression (7), i.e. $P_1 = \mu + \frac{\rho}{\tau_\theta}u_1$.

2.2 Equilibrium prices at $t = 0$

Now we turn to the portfolio choices of investors at $t = 0$ and to the resulting equilibrium price P_0. At $t = 0$ only the arbitrageur has private information, all other investors being still uninformed. Hence investors' portfolio choice problem is similar to that in (2), with the difference that it is not conditioned on information Ω:

[7] In Section 5 we will relax this assumption and consider the case of arbitrageurs whose trades affect prices.

$$\max_{\{y_i\}_{i \in N}} \sum_{i=1}^{N} \left[y_{i0}(\mathbb{E}(P_{i1}) - P_{i0}) - \frac{\rho}{2} y_{i0}^2 \mathbb{V}(P_{i1}) \right]. \tag{8}$$

As of $t = 0$, the distribution of the payoff θ of a representative asset (dropping the index i for simplicity) is $N(\mu, \sigma_\theta^2)$ for all investors. As the probability of the arbitrageur advertising any given asset is negligible, from the standpoint of an investor at $t = 0$ the price at $t = 1$ is

$$P_1 = \mu + \frac{\rho}{\tau_\theta} u_1,$$

which is distributed according to $N(\mu, \frac{\rho^2}{\tau_\theta^2}\sigma_u^2)$. For each asset, investors maximize the certainty equivalent of its payoff:

$$\max_{\{y_0\}} (\mu - P_0) y_0 - \frac{\rho^3}{2\tau_\theta^2} \sigma_u^2 y_0^2, \tag{9}$$

so that their equilibrium demand at $t = 0$ is

$$y_0 = \frac{\mu - P_0}{\frac{\rho^3}{\tau_\theta^2}\sigma_u^2} \tag{10}$$

Since noise traders buy a random amount $u_0 \sim N(0, \sigma_0^2)$ and the arbitrageur is atomistic, i.e. his trades at $t = 0$ have no impact on P_0, market clearing requires $y_0 + u_0 = 0$, so that the equilibrium price at $t = 0$ is

$$P_0 = \mu + \frac{\rho^3}{\tau_\theta^2} \sigma_u^2 u_0, \tag{11}$$

i.e. it equals the expected fundamental value plus a term reflecting the price impact of noise traders' purchases.

3 The Arbitrageur's Strategy

The arbitrageur has two interrelated decisions to take: portfolio choice and advertising. To find his optimal strategy, we proceed in three steps. First, we analyze his investment in advertised assets at $t = 0$, taking his decision to advertise them as given. Second, we find his optimal advertising decision, by identifying the assets that he advertises at $t = 1$.

Finally, we characterize his optimal portfolio under optimal advertising.

3.1 Investment decision

As of $t = 0$, the arbitrageur has private signals $\hat{\theta}_i$ about assets $i \in \mathbf{M}$. His expected utility maximization is similar to the mean-variance optimization problem faced by other investors in (2):

$$\max_{\{x_i\}_{i\in N}} \sum_{i\in\mathbf{M}} \left[x_i(\mathbb{E}(P_{i1}|\hat{\theta}_i) - P_{i0}) - \frac{\rho}{2}x_i^2\mathbb{V}(P_{i1}|\hat{\theta}_i)\right] + \sum_{i\in\mathbf{N}\setminus\mathbf{M}} \left[x_i(\mathbb{E}(P_{i1}) - P_{i0}) - \frac{\rho}{2}x_i^2\mathbb{V}(P_{i1})\right]. \tag{12}$$

We first solve this optimization problem for the assets that the arbitrageur is informed about, taking for the time being as given the advertising decision, and therefore also the mass m of investors who process the advertisement about the asset. As of $t = 0$, the arbitrageur expects the price P_1 in (7) to be distributed according to $P_1 \sim N(\hat{P}, \sigma_P^2(m))$, where

$$\hat{P} = \mathbb{E}[P_1|\hat{\theta}] = \frac{m\tau_A\hat{\theta} + \tau_\theta\mu}{m\tau_A + \tau_\theta}, \tag{13}$$

and

$$\sigma_P^2(m) = \mathbb{V}[P_1|\hat{\theta}] = \frac{\rho^2\sigma_u^2}{(\tau_\theta + m\tau_A)^2}. \tag{14}$$

As the arbitrageur will liquidate his position x in all assets at $t = 1$, his returns are normal and independently distributed, and his portfolio problem (12) is equivalent to choosing a position x in any asset so as to maximize the certainty-equivalent profit:

$$\max_{\{x\}} \; (\hat{P} - P_0)x - \frac{\rho}{2}x^2\sigma_P^2(m). \tag{15}$$

Substituting for P_0 and $\hat{P}$ from (11) and (13) we can express the arbitrageur's optimal investment as

$$x = \frac{\hat{P} - P_0}{\rho\sigma_P^2(m)} = \left(m\tau_A\tau_\theta^2(\hat{\theta} - \mu) - \rho^3\sigma_u^2(\tau_\theta + m\tau_A)u_0\right)\frac{\tau_\theta + m\tau_A}{\rho^3\sigma_u^2\tau_\theta^2}. \tag{16}$$

Replacing x from this expression in (15) yields the arbitrageur's expected payoff from investing in an asset:

$$\pi(m) = \frac{(\hat{P} - P_0)^2}{2\rho} \frac{(\tau_\theta + m\tau_A)^2}{\rho^2 \sigma_u^2} = \left(m\tau_A(\hat{\theta} - \mu) - \frac{\rho^3}{\tau_\theta^2} \sigma_u^2 (\tau_\theta + m\tau_A) u_0 \right)^2 \frac{1}{2\rho^3 \sigma_u^2}. \quad (17)$$

This expression reveals that there are two sources of profit for the arbitrageur. First, he can exploit his private information, his potential informational rent being captured by the term $m\tau_A(\hat{\theta} - \mu)$. Second, like any other speculator, he can exploit the price impact of noise trading shocks at $t = 0$ by taking offsetting positions to their orders u_0: the resulting profit is captured by the term $\frac{\rho^3}{\tau_\theta^2}\sigma_u^2(\tau_\theta + m\tau_A)u_0$.

These two trading motives can either complement each other or not. Consider for instance the case in which the arbitrageur receives a positive private signal about the asset's value ($\hat{\theta} - \mu > 0$): then, if noise traders happen to sell the asset ($u_0 < 0$) and thus depress the price at $t = 0$, the two trading motives will reinforce each other in pushing the arbitrageur to buy the asset ($x > 0$). If instead noise traders happen to buy the asset ($u_0 > 0$) and thus raise the price at $t = 0$, the two trading motives conflict with each other. In the latter case, if noise traders' purchases are large enough, the second trading motive can dominate, so that the arbitrageur will sell the asset, trading against his private signal. If so, the arbitrageur will solely want exploit noise traders at $t = 0$ as a classic speculator and not advertise his private signal at all.

The arbitrageur will want to advertise his information only if the two trading motives complement each other or if the informational advantage is the main source of his potential profits. In this case, he will want to broadcast his ad as widely as possible: the larger the fraction of investors m that he manages to inform, the greater his informational rent per dollar invested in the asset, and therefore the larger the sum that he initially wishes to invest. On the other hand, the larger his investment, the greater is his incentive to advertise the asset widely. As we shall see, this complementarity between advertising and investment is a key feature of the model. This intuition explains the following result:

Proposition 1 (No advertising case). *The arbitrageur's expected payoff from an asset is convex in the fraction m of investors informed about the asset, so that he prefers either not to advertise the asset at all ($m = 0$) or to inform all investors about it ($m = 1$). The arbitrageur will not advertise assets for which his expected payoff from exploiting only noise* *trading at* $t = 0$ *exceeds the expected payoff from advertising* $(\pi(0) \geq \pi(1))$, *i.e.:*

$$\rho^3 \sigma_u^2 \tau_\theta |u_0| \geq |\tau_A \tau_\theta^2 (\hat{\theta} - \mu) - \rho^3 \sigma_u^2 (\tau_\theta + \tau_A) u_0|. \tag{18}$$

The proof is immediate, once it is recognized that $\pi(m)$ is a convex function, its second derivative being positive:

$$\frac{\partial^2 \pi(m)}{\partial m^2} = \left(\tau_A (\hat{\theta} - \mu) - \frac{\rho^3}{\tau_\theta^2} \sigma_u^2 \tau_A u_0 \right)^2 \frac{1}{\rho^3 \sigma_u^2} > 0.$$

As can be seen from (18), the arbitrageur prefers not to advertise assets for which the noise trade shock at $t = 0$ is not only large in absolute value (large $|u_0|$), but also pushes the price P_0 in the same direction as his private information would recommend. In this case, by chance noise traders at $t = 0$ already move the price towards its fundamental value, reducing the profitability of advertising for the arbitrageur, so that, rather than advertising the asset, the arbitrageur will simply exploit noise trading at $t = 0$.

Assets for which condition (18) holds will not be advertised. For these assets, the arbitrageur's profits will stem solely from trading against the order flow coming from noise traders, just as for any other uninformed investor at $t = 0$. To see this, note that for these assets the arbitrageur's expected price at $t = 1$ equals μ (as can be seen by setting $m = 0$ in expression (13)), so that his optimal investment is the same as that of uninformed investors given by (10).[8] Clearly, the same applies to all assets the arbitrageur has no private information about at $t = 0$, which he trades as other uninformed investors.

Since the distinctive feature of advertised assets is the arbitrageur's ability to profit from his private information about them by impounding it in their price at $t = 1$, in what follows we shall disregard the profits from noise trading shocks at $t = 0$, u_0, assuming them to be negligible relative to the arbitrageur's rents from private information $\tau_A(\hat{\theta} - \mu)$, that is:

$$\frac{\rho^3 \sigma_u^2 (\tau_\theta + \tau_A) u_0}{\tau_A \tau_\theta^2 (\hat{\theta} - \mu)} \approx 0, \tag{19}$$

[8]Note that the arbitrageur cannot profit from his superior information about the fundamental value of an asset by simply trading on it. The reason is that this information will only become publicly known as of $t = 2$, while the arbitrageur has to liquidate his portfolio at $t = 1$. In other words, his short holding period implies that he can only profit from his long-term information by advertising it.

so that condition (18) does not hold. This is warranted not only when the realization of noise trading shock is small in absolute size $|u_0|$, but also when the variance of u_1 (i.e., σ_u^2) and risk aversion ρ are small, the precision of initial information τ_θ is high, and most importantly the arbitrageur's private valuation of the asset greatly differs from the market's prior, i.e. $|\hat{\theta} - \mu|$ is large. Using (19), the arbitrageur's expected payoff (17) simplifies to:

$$\pi(m) = \frac{m^2 \tau_A^2 (\hat{\theta} - \mu)^2}{2\rho^3 \sigma_u^2}. \tag{20}$$

From this expression it is immediate that the arbitrageur's expected gain is increasing and convex in the fraction m of investors reached by his advertisement, in the initial mispricing $\hat{\theta} - \mu$ and in the precision of the arbitrageur's information τ_A. The expression also implies that the arbitrageur is harmed by the variance of noise trading σ_u^2, since this creates liquidation risk at $t = 1$ for him.[9]

3.2 Advertising decision

We now consider the advertising decision taken by the arbitrageur. We start our analysis from the case where the assets for which the arbitrageur has superior information are symmetric, in the sense that their returns are identically and independently distributed and that learning about any of them requires the same attention from investors. Next, we turn to the more general case where these assets may differ both in the distribution of their returns and in the attention that they require from investors.

Symmetric assets. If the arbitrageur advertises L assets, then from (1) the mass of investors that pay attention to his advertising about each of them is $m(L) = \min[1, K/L]$. The arbitrageur's expected payoff from an optimal position in any advertised asset $i \in \mathbf{L}$, denoted by π_{Ai}, is given by expression (20), while his expected payoff from an optimal position in a non-advertised asset $j \in \mathbf{M} \setminus \mathbf{L}$, denoted by $\pi_{Nj} = 0$, can be obtained by setting $m = 0$ in (20). The optimal number of advertised assets, $L \leq M$, maximizes his expected payoff:

[9] However, this result does not necessarily hold if (19) is violated, as shown by equation (17): intuitively, the variance of noise trading σ_u^2 increases the speculative profit from trading against noise traders at $t = 0$ as well as the arbitrageur's liquidation risk at $t = 1$, so that its overall effect on the arbitrageur's payoff is ambiguous. We thank an anonymous referee of this paper for bringing this result to our attention.

$$\Pi(L) = \sum_{i \in L} \pi_{Ai} + \sum_{j \in M \setminus L} \pi_{Nj} = L\,(m(L))^2\,\frac{\tau_A^2(\hat{\theta} - \mu)^2}{2\rho^3\sigma_u^2}, \tag{21}$$

which yields the following result:

Proposition 2 (Concentrated advertising). *The arbitrageur advertises* $\min(K, M)$ *assets, so that for each of them the fraction of informed investors is* $m = 1$.

Hence, if the arbitrageur is informed about at least as many assets as an individual investor can learn about, i.e., $M \geq K$, he will advertise exactly K assets. The reason why he will advertise no fewer than K assets is intuitive: leaving investors' attention capacity unexploited would be wasteful, as the arbitrageur's information would not be impounded in market prices for some assets. What is less intuitive is why the arbitrageur does not want to advertise more than K assets: after all, investors can always disregard the information that they cannot digest! However, the arbitrageur has the incentive not to over-exploit investors' attention. To understand why, note that if he advertised more than K assets, each ad would compete for investors' attention against the others. As attention is split evenly among ads, this would lower the mass of informed traders paying attention to each ad, and so moderate the price correction induced by it. Conversely, the more numerous are the investors who pay attention to an ad, the closer the price of the corresponding asset at $t = 1$ will be to the arbitrageur's estimate $\hat{\theta}$ of its fundamental value: this will reduce the riskiness of the arbitrageur's position in the asset, and increase his gain from investing in it at $t = 0$. In turn, this prompts the arbitrageur to take a larger position in the asset than in any non-advertised asset.

The complementarity between investment and advertising already established in Proposition 1 is also at the root of the result in Proposition 2: as the arbitrageur wishes to make his information-related positions as safe as possible, he wants to restrict the number of assets to be advertised to those that investors can fully learn about, given their attention capacity. Indeed, anecdotal evidence suggests that investors' attention span is limited to a few stocks at a time: for instance, in conferences where hedge funds pitch stocks to institutional investors, each fund typically provides information about a single stock only.

Clearly, these results depend on the assumption that investors have limited attention capacity: if their capacity were unlimited ($K \to \infty$), the arbitrageur would advertise all assets he is informed about. The friction created by limited attention capacity could alternatively be modeled by assuming that investors face costs in processing the ads issued by the arbitrageur: if these costs were increasing in the number of processed ads, the arbitrageur would again concentrate his advertising activity on a limited set of assets. The same prediction would obtain if the friction were to arise from the advertising technology itself: it can be shown that, if the probability of an ad being processed by an investor were an increasing and concave function of a costly effort incurred by the arbitrageur, the latter would optimally advertise a single asset. Hence, the result of concentrated advertising in Proposition 2 is quite robust to changes in the way the friction in the communication between the arbitrageur and investors is modeled.

Heterogeneous assets. If assets are symmetric, as assumed above, then the arbitrageur is indifferent as to which ones to advertise; this is no longer the case if instead they are heterogeneous. Assets may differ in terms of the conditional distribution of their return from the arbitrageur's viewpoint, namely, their expected return $\hat{\theta}_i - \mu_i$, the precision of the prior information τ_{θ^i} and of the arbitrageur's signal τ_{Ai}, and the variance of noise trading $\sigma^2_{u_i}$. But they can also differ in the difficulty D_i that investors face in processing the arbitrageurs' information, in which case the investors' attention constraint becomes $\sum_{i \in \mathbf{K}} D_i \leq C$, where $\mathbf{K}$ is the set of arbitrageur's ads that an investor can process. For simplicity, we assume ads to feature one of two possible levels of difficulty: those containing simple information present a low level of difficulty $\underline{D} > 0$ for investors, while those containing complex information feature a high level of difficulty $\overline{D} > \underline{D}$. For brevity, we refer to the two asset types as "simple" and "complex", respectively.

Recall that, from Proposition 2, if the arbitrageur advertises $L = K$ symmetric assets, then he does not wish to advertise an additional one: he fully exploit each investor's attention capacity, but does not exceed it, so that $LD \leq C$. A similar logic applies when ads about assets require different levels of attention: also in this case, processing the signals about the advertised assets should not exceed each investor's information processing capacity ($\sum_{i \in \mathbf{L}} D_i \leq C$), but in this case the optimal number of assets also depends on the cross-sectional distribution of their characteristics and on the demands they place on investors' attention.

To see this, first suppose that the set of ads issued by the arbitrageur does not exhaust each investor's information processing capacity, i.e., $\sum_{i\in\mathbf{L}} D_i \leq C$. In this case, all investors can process these ads, so that the arbitrageur's expected payoff is found by setting $m = 1$ in expression (20):

$$\pi_{Ai} = \frac{\tau_{Ai}^2(\hat{\theta}_i - \mu_i)^2}{2\rho^3\sigma_{u_i}^2}, \tag{22}$$

whereas not advertising asset i would yield a zero expected payoff.

Clearly, the arbitrageur prefers to advertise assets that deliver high expected payoff π_{Ai}, but must also take into account the attention capacity "consumed" by the corresponding ads, hence their level of difficulty. To capture this trade-off, we characterize each asset i in the set $\mathbf{M}$ by its expected payoff per unit of required attention, i.e. scaled by the difficulty of processing the corresponding ad:

$$\frac{\pi_{Ai}}{D_i} = \frac{\tau_{Ai}^2(\hat{\theta}_i - \mu_i)^2}{2\rho^3\sigma_{u_i}^2 D_i}. \tag{23}$$

The arbitrageur's total expected payoff is maximized by advertising the assets that feature the highest values of the ratio (23). Hence, in selecting the L^* assets to be advertised, the arbitrageur will rely on the following algorithm. First, he will rank the assets in $\mathbf{M}$ by decreasing values of their payoff per unit of required attention as defined by (23), so that $\pi_{Ai-1}/D_{i-1} > \pi_{Ai}/D_i$, $i = 2, .., M$. His total expected payoff will increase with the addition of each of these assets, until he hits the investors' attention capacity constraint. At this point, two situations can occur. If the next asset i to be added according to criterion (23) is a simple one (i.e., $D_i = \underline{D}$), then the arbitrageur wishes to advertise only the already selected assets, as the additional one would violate investors' capacity constraint. If instead the next asset i according to (23) is a complex one ($D_i = \overline{D}$), then the arbitrageur might still be able to add one or more simple assets (featuring difficulty $\underline{D}$) and still not exceed the investors' attention capacity constraint. Clearly, when adding simple assets he will pick those with the highest benefit-cost ratio (23). Eventually, he will reach the point when no further simple asset can be added without violating this constraint.

While this may seem an obvious selection algorithm, establishing its optimality is not trivial, because in principle the arbitrageur could choose to advertise more assets than any individual investor can process, and let investors randomly allocate their attention to advertised assets. However, we establish that this is never optimal:

Proposition 3 (**Concentrated advertising with heterogeneous assets**). *The arbitrageur never advertises assets that in total require more attention than an investor's individual capacity, i.e. he chooses to advertise assets* $\mathbf{L}^*$ *such that* $\sum_{i\in\mathbf{L}^*} D_i \leq C$.

The selection algorithm based on the benefit-cost ratio (23), together with Proposition 3, characterizes the set of assets that the arbitrageur chooses to advertise:

Proposition 4 (**Characteristics of advertised assets**). *The expected payoff per unit of attention from advertising asset* i *is increasing in its mispricing* $(\hat{\theta}_i - \mu_i)^2$ *and in the precision of the arbitrageur's private information* τ_{Ai}, *and is decreasing in the variance* $\sigma^2_{u_i}$ *of noise trading and in the difficulty* D_i *of the information about the asset.*

These comparative statics are intuitive: the arbitrageur will seek to advertise the assets with the highest expected return and lowest risk, while taking into account the required attention. His expected payoff is highest when mispricing $(\hat{\theta}_i - \mu_i)^2$ is greatest and, for given mispricing, his risk is lowest for the assets he is best informed about (high τ_{Ai}) and those that entail the least noise trading risk ($\sigma^2_{u_i}$ low).[10] Moreover, other things being equal, assets for which arbitrageurs' private information is simple (low D_i) are more likely to be advertised than those for which their information is complex, the former being easier to digest for investors than the latter.

3.3 Optimal portfolio

We can now fully characterize the optimal portfolio of the arbitrageur at $t = 0$. We assume $M \geq K$. As argued in the previous section, he chooses to advertise K assets, whose characteristics are illustrated by Proposition 4. Then, his optimal investment in the advertised asset i is obtained by setting $m = 1$ in expression (16) and using condition (19):

$$x_i = \frac{\tau_{Ai}(\hat{\theta}_i - \mu_i)}{\rho^3 \sigma^2_{u_i}}, \tag{24}$$

[10] By the same token, the arbitrageur would not benefit from adding noise to the private information communicated to the market, unlike the informed investors in Bommel (2003), Brunnermeier (2005) and Indjejikian et al. (2014). In these models, the informed investor benefits from the noise his announcement injects in market prices, and thus hinders price discovery; in contrast, our arbitrageur benefits from reducing noise in the market price, and accordingly his advertising aims to enhance price discovery.

while his investment in non-advertised assets j is zero: $x_j = 0$. The arbitrageur takes positions only in the K assets that he advertises: he will not trade other assets, even if he has superior information about them, because such information refers to their final value at $t = 2$, while he must liquidate his position at $t = 1$. Hence taking a position in non-advertised assets would expose the arbitrageur to noise trading risk at $t = 1$, without generating any offsetting profits at the expense of noise traders at $t = 0$ (by our simplifying assumption that $u_0 = 0$). By the same token, the arbitrageur does not take any position in assets he is not informed about as of $t = 0$.

Expression (24) shows that the arbitrageur's positions in advertised assets depend on the same characteristics that induce him to advertise them in the first place, the intuition being the same as that behind Proposition 4:

Proposition 5 (Arbitrageur's portfolio). *The arbitrageur's position in an advertised asset i is increasing in its expected appreciation $\hat{\theta}_i - \mu_i$, in the precision of the arbitrageur's private information τ_{Ai}, and is decreasing in the variance $\sigma^2_{u_i}$ of noise trading at $t = 1$.*

4 Multiple Arbitrageurs

Thus far we have considered a setting with a single informed arbitrageur, who monopolizes investors' attention. If multiple arbitrageurs compete for attention, the total number of assets advertised is no longer set by a single arbitrageur: it is the result of all arbitrageurs' advertising choices. This section shows that the outcome can differ substantially from that of the monopolist, unless arbitrageurs coordinate their actions.

If there are A arbitrageurs, each informed about M different assets, the total number of assets that can be advertised is $A \cdot M$. If each investor has the capacity to process information about all these assets, i.e. $A \cdot M \leq K$, then trivially arbitrageurs will advertise all the private information they have: in this case, the limit to information processing capacity is not binding. Clearly, the more interesting case is that in which investors cannot process all the signals, i.e. $A \cdot M > K$. In principle, arbitrageurs could collectively end up advertising all $A \cdot M$ assets for which they have information, but as we shall see, this would lead them to over-exploit the common resource of investors' attention, as in the tragedy of the commons: advertising too many assets would lead to too little attention being paid to each asset and therefore excessive persistence of mispricing.

To mitigate this inefficiency, arbitrageurs need to coordinate. Whether they can do so depends partly on the commonality of their information. They may have "common information" about the same set of assets or "exclusive information" about distinct sets of assets. Common information may arise either because they happen to receive the same signal independently, or because they share information prior to advertising and trading. The common information case is the more interesting of the two, as it is the one where arbitrageurs may be able to solve the coordination problem: when they have superior information about the same assets, they may informally agree on which ones to advertise. Hence, we focus the analysis mainly on this case, leaving the discussion of exclusive information to Section 4.3.

In most of the analysis, for simplicity we consider identical assets with independent returns, but briefly illustrate how the results would change with heterogeneous assets. This extension will highlight another interesting margin along which inefficiencies may arise in advertising: lacking coordination, even arbitrageurs with common information may end up advertising the "wrong" assets, for instance those featuring mild rather than severe mispricing.

4.1 Coordination with common information

Consider first the benchmark case where arbitrageurs have information about the same set of $A \cdot M$ assets, and may agree to advertise $L \leq A \cdot M$ of them. Each investor randomly picks K assets to learn about from the set of advertised assets, so that the fraction of investors who are informed about any advertised asset is

$$m(L) = \min\left[1, K/L\right]. \tag{25}$$

Each arbitrageur i knows that L assets are being advertised, so that his total expected payoff $\Pi(L)$ is given by expression (21). Because all arbitrageurs have the same information and coordinate, they will choose L so as to maximize $\Pi(L)$. Hence, Proposition 2 applies: arbitrageurs will just saturate the information processing capacity of investors by setting $L = K$, so that $m = 1$.

Hence, if arbitrageurs coordinate their decisions, the concentrated advertising principle of Proposition 2 carries over to multiple arbitrageurs. Notice that in this case the prediction is not that each and every arbitrageur should advertise the same set of assets, but rather that if some assets are advertised by any of them, other arbitrageurs should not distract investors by advertising other assets. Just as a single arbitrageur does not want to advertise several assets, in order to avoid "dispersing" investors' attention, multiple arbitrageurs will refrain from advertising assets that differ from those advertised by others: each has the incentive to "piggyback" on the others' advertising activity. Whether advertising is done by a single arbitrageur or by a coordinated group, the complementarity between advertising and investment decisions is at the core of the concentration result.[11] Even if assets were heterogeneous, the choices of coordinated arbitrageurs would be identical to those of a monopolistic arbitrageur: in both cases, they will advertise the most severely mispriced assets, by Proposition 4.

This equilibrium behavior may appear to resemble the herding induced by information cascades, but in fact it is quite different: in this model, the fact that all arbitrageurs pick the same assets depends on common fundamental information and strategic complementarity, not on an attempt to gather useful information from others' decisions. Indeed, their correlated behavior speeds up price discovery, rather than delaying it as in models of cascades.

4.2 No coordination with common information

Now we turn to the more interesting case where arbitrageurs have common information about assets but fail to coordinate on which ones to advertise: when each of them chooses his portfolio and trades at $t = 0$, he does not know yet which assets will be advertised by the others. Let us redefine by $\mathbf{M}$ the set of $A \cdot M$ assets that arbitrageurs can collectively advertise, of which an arbitrageur can choose to advertise any subset $\mathbf{m}_i \subseteq \mathbf{M}$.

Assuming that no attention is required to recognize identical messages, each investor will pay attention to only one ad per asset. Since each investor can process at most K ads, he will pick them randomly from the L unique ads in the set $\mathbf{L} = \bigcup_{j \in A} \mathbf{m}_j = \mathbf{L}_i \cup \mathbf{L}_{-i}$,

[11] This result would also hold if advertising were costly: in this case arbitrageurs would have an additional reason to avoid advertising different assets, namely, avoiding duplication of advertising costs.

where $\mathbf{L}_i$ includes the L_i assets advertised *only* by arbitrageur i, and $\mathbf{L}_{-i}$ the L_{-i} assets advertised by other arbitrageurs. The total number of unique ads is $L_i + L_{-i}$, and the fraction of investors who are informed about any advertised asset is

$$m(L_i, L_{-i}) = \min\left[1, \frac{K}{L_i + L_{-i}}\right]. \tag{26}$$

To find how many assets are advertised in the absence of coordination, we need to identify the best response L_i of arbitrageur i to the number L_{-i} of assets advertised by others. By expression (21) the expected payoff of arbitrageur i is

$$\Pi(L_i, L_{-i}) = (L_i + L_{-i})\frac{\Big(m(L_i, L_{-i})\tau_A(\hat{\theta} - \mu)\Big)^2}{2\rho^3\sigma_u^2}. \tag{27}$$

Two cases can arise. First, suppose that other arbitrageurs advertise $L_{-i} < K$ assets. If the arbitrageur i chooses to advertise $L_i \leq K - L_{-i}$ assets, then $m = 1$ and his expected payoff is increasing in L_i, up to $L_i = K - L_{-i}$:

$$\Pi(L_i, L_{-i}) = (L_i + L_{-i})\frac{\tau_A^2(\hat{\theta} - \mu)^2}{2\rho^3\sigma_u^2}.$$

If the arbitrageur were to advertise $L_i \geq K - L_{-i}$ his expected payoff payoff would be

$$\Pi(L_i, L_{-i}) = \frac{K^2}{L_i + L_{-i}}\frac{\tau_A^2(\hat{\theta} - \mu)^2}{2\rho^3\sigma_u^2}, \tag{28}$$

which is decreasing in L_i. Hence, arbitrageur i's best response is to advertise $L_i = K - L_{-i}$ additional assets, so as to precisely saturate investors' informational capacity. This implies that at least K assets are advertised in equilibrium.

Alternatively, suppose other arbitrageurs advertise $L_{-i} \geq K$ assets. Then the expected payoff of arbitrageur i is given by expression (28), and his best response is not to advertise any additional asset, i.e. set $L_i = 0$. Note that arbitrageur i can still optimally choose to advertise the same L_{-i} assets already advertised by others, but he would not find it optimal to advertise any other asset, as this would lower his payoff (28).

This shows that, when arbitrageurs do not coordinate, there are multiple equilibria, in some of which the number of advertised assets exceeds investors' processing capacity, i.e.

$L_{-i} > K$, as stated in the next proposition, where we consider symmetric equilibria in which all arbitrageurs advertise the same assets, i.e. $L_i = L$ for any i.

Proposition 6 (Multiple equilibria without coordination). *With common information and no coordination, there are multiple equilibria: any number $L \in [K, A \cdot M]$ of assets can be advertised in equilibrium.*

The proof follows immediately from the arguments mentioned above. First, at least K assets are advertised in equilibrium ($L \geq K$). Second, in a symmetric equilibrium each arbitrageur advertises the same assets ($L_i = L$), and no arbitrageur wants to deviate. The equilibrium payoff of the arbitrageur in an equilibrium with $L \geq K$ advertised assets is

$$\Pi^*(L) = \frac{K^2}{L} \frac{\tau_A^2(\hat{\theta} - \mu)^2}{2\rho^3\sigma_u^2},$$

which is highest for $L = K$. Hence, equilibria with $L > K$ are inefficient, as too many assets are advertised, leading to an excessive dispersion of investors' attention from the viewpoint of arbitrageurs.

If the assets about which arbitrageurs are informed differ in their characteristics, competition entails another possible inefficiency, namely that the advertised assets are not those yielding the highest possible payoff to arbitrageurs. To illustrate this point, suppose that there are only two types of assets, a "good" asset (G) and a "bad" one (B), such that advertising asset G yields a greater gain for arbitrageurs than advertising asset B, for instance because the former is more mispriced than the latter:

$$\pi_G \equiv \frac{\tau_{AG}^2(\hat{\theta}_G - \mu_G)^2}{2\rho^3\sigma_{uG}^2} \geq \pi_B \equiv \frac{\tau_{AB}^2(\hat{\theta}_B - \mu_B)^2}{2\rho^3\sigma_{uB}^2}. \tag{29}$$

Suppose also that half of the assets about which the arbitrageur is informed are good and the other half bad, and the number of assets in each class is sufficient to saturate investors' capacity ($A \cdot M/2 > K$). Clearly, a monopolistic arbitrageur would always prefer to advertise the good rather than the bad assets. However, this may not be the case when multiple arbitrageurs do not coordinate in deciding which assets to advertise: asset heterogeneity may entail an additional source of inefficiency with multiple arbitrageurs.

Indeed, if assets are not too different, i.e. if the following condition holds

$$\pi_G \leq \pi_B \left(2 + \frac{1}{AM}\right), \tag{30}$$

then there is an equilibrium where only $L \geq K$ bad assets are advertised, another where only good ones are advertised, and others with any combination of the two. More generally, we show that:

Proposition 7 (Inefficient advertising with common information). *With common information and no coordination, if condition (30) holds, then in equilibrium any combination of good and bad assets can be advertised, the total number of advertised assets being* $L \in [K, A \cdot M]$.

Intuitively, condition (30) guarantees that even if only bad assets are advertised by arbitrageurs, none of them wants to advertise a good one. Hence, bad assets can be advertised in equilibrium even if all arbitrageurs are aware that they could advertise assets that deliver a larger expected payoff. This stems from the strategic complementarity between the advertising and investment decisions, which makes arbitrageurs' payoff from an asset convex in the fraction of investors paying attention to it. If for instance assets with only mild mispricing are advertised, even advertising a more severely mispriced one would reduce the fraction of investors informed about each and, due to convexity, disproportionately reduce arbitrageurs' payoff.

This inefficiency is due to lack of coordination among arbitrageurs. In fact, it did not arise in Section 4.1 where arbitrageurs were assumed to coordinate: in that setting, they would collectively choose to advertise the best assets only, e.g., those with the largest mispricing $(\hat{\theta} - \mu)^2$. This result may explain why financial markets sometimes focus on minor mispricing of some assets while neglecting much more significant mispricing of others, such as RMBSs or CDOs before the recent financial crises. Our theory provides a new explanation for the persistence of substantial mispricing, which differs from those already proposed in the literature on limits to arbitrage, where mispricing persists because arbitrageurs have limited resources (Shleifer and Vishny (1997)) or are deterred by noise-trader risk (DeLong et al. (1990)) or synchronization risk (Abreu and Brunnermeier (2002)). In contrast to these explanations, in our setting arbitrageurs would have the resources and the ability to eliminate substantial mispricing, if only they could coordinate their investment and advertise the most mispriced assets.

This also predicts trading to be strongly correlated across arbitrageurs, as in equilibrium they have the incentive to trade advertised assets more intensively than other assets they are informed about. This is consistent with evidence by Luo (2018) that around the date in which a hedge fund pitches a stock at an investment conference, other hedge funds take similar positions, and they all liquidate these positions subsequently, like the arbitrageurs in our model. These correlated trading strategies are reminiscent of the "wolf pack" activism by hedge funds documented by Becht et al. (2017) and modeled by Brav et al. (2021). Just as in their model activists implicitly coordinate with many followers in engaging target management, in our equilibria informed arbitrageurs trade the same advertised assets, resulting in highly correlated trading even though they act non-cooperatively. Of course, correlated trading emerges *a fortiori* if arbitrageurs can coordinate, as assumed in Section 4.1. Yet such coordination is not necessary for them to follow correlated strategies, as these are also part of the non-cooperative equilibria derived in this section.

4.3 Exclusive information

Finally, let us briefly consider the case in which each arbitrageur has an exclusive information advantage about M different assets, and decides on advertising independently from others, so that there is neither common information nor coordination. Also in this case, we denote the number of assets advertised by arbitrageur i by L_i and the number of those advertised by other arbitrageurs by L_{-i}. Since now by definition arbitrageurs never advertise the same assets, the fraction of informed investors about any advertised asset is given by expression (26).

As information is exclusive, no arbitrageur knows which assets others are going to advertise, but they all know that each of them will advertise a different set of assets. Hence, each arbitrageur i will only invest in the L_i assets that he will advertise, so that from (21) his expected payoff is

$$\Pi(L_i, L_{-i}) = L_i \frac{\left(m(L_i, L_{-i})\tau_A(\hat{\theta} - \mu)\right)^2}{2\rho^3\sigma_u^2}. \tag{31}$$

Unlike in the common information case, here no benefit accrues to the arbitrageur from the advertising activity of other arbitrageurs. As a result, equilibrium advertising with exclusive information also differs from those obtained under common information, and indeed invariably leads to a lower payoff, as arbitrageurs always overload investors with advertising:

Proposition 8 (Inefficient advertising with exclusive information). *If there are more than two arbitrageurs, then in an equilibrium with exclusive information each of them advertises all M assets he is informed about, so that the total number of advertised assets $A \cdot M$ inefficiently exceeds investors' information processing capacity K.*

Effectively, advertising with exclusive information results in the most inefficient arrangement: each arbitrageur overloads investors' attention to the maximum possible extent, as he only cares about a different subset of assets, and therefore does not take into account the cost of diluting investors' attention about other assets. Hence, it is the polar opposite arrangement relative to common information with coordinated advertising, which maximizes the joint payoff of arbitrageurs, while common information without coordination delivers an intermediate expected payoff to arbitrageurs relative to these two extreme cases.

5 Price Impact

Throughout the foregoing analysis, the trades of arbitrageurs were assumed to be small relative to the market, and thus to have no price impact. It is natural to ask how the results are affected if arbitrageurs' trades have price impact, both at $t = 0$ upon building their initial position and upon divesting at $t = 1$. In both cases, price impact will reduce the profitability of arbitrage: for instance, an arbitrageur who buys upon receiving good news about an asset will tend to raise its price, and symmetrically will lower its resale price when he liquidates it. In this section we explore how these two adverse price impacts affect the arbitrageur's optimal advertising strategy.

For analytical tractability, we revert to the case of a monopolistic arbitrageur and symmetric asset returns, and thus drop asset-specific subscripts in the analysis. We also retain the assumption (made throughout the analysis) that at $t = 0$ investors are unaware of where the arbitrageur's informational advantage lies, and therefore do not learn from prices. Hence, the price pressure from the arbitrageur's trades does not arise from investors' inference about the asset value upon observing the net order flow but from the arbitrageur's order flow impacting a price-elastic demand by other investors.

The arbitrageur will scale back his trades to reduce their adverse price impact on the market-clearing price at $t = 0$ and, symmetrically, on the expected market-clearing price at $t = 1$. Specifically, the arbitrageur takes into account that at $t = 1$ the price will balance the informed investors' trade my_I, the order $(1 - m)y_U$ placed by the fraction $1 - m$ of rational investors who remain uninformed, the noise trade u_1 and arbitrageur's liquidation of his initial investment x. As rational investors are atomistic, their individual trades have no price impact, and are given by the same formulas as in Section 2. Hence, the market clearing condition is

$$(1 - m)y_U + my_I + u_1 - x = 0,$$

and the resulting price is similar to expression (7):

$$P_1 = \frac{m\tau_A\hat{\theta} + \tau_\theta\mu}{m\tau_A + \tau_\theta} + \frac{\rho}{m\tau_A + \tau_\theta}(u_1 - x). \tag{32}$$

Based on this expression, at $t = 0$ the arbitrageur will expect the price at $t = 1$ to be distributed according to $P_1 \sim N(\hat{P}_1, \sigma_P^2(m))$, where

$$\hat{P}_1 = \mathbb{E}[P_1|\hat{\theta}] = \frac{m\tau_A\hat{\theta} + \tau_\theta\mu}{m\tau_A + \tau_\theta} - \frac{\rho}{m\tau_A + \tau_\theta}x, \tag{33}$$

and

$$\mathbb{V}[P_1|\hat{\theta}] = \sigma_P^2(m) = \frac{\rho^2\sigma_u^2}{(m\tau_A + \tau_\theta)^2}. \tag{34}$$

Similarly, at $t = 0$ the price will balance the demand of rational investors and noise traders with the arbitrageur's initial investment. The demand by investors is the same as in (10), namely, $y_0 = (\mu - P_0)/(\rho^3\sigma_u^2/\tau_\theta^2)$, the order placed by noise traders is u_0 and that placed by the arbitrageur is x:

$$\frac{\mu - P_0}{\rho^3\sigma_u^2/\tau_\theta^2} + u_0 + x = 0, \tag{35}$$

so that the equilibrium price at $t=0$ is

$$P_0 = \mu + \rho^3 \frac{\sigma_u^2}{\tau_\theta^2}\left(u_0 + x\right). \tag{36}$$

Hence, at $t=0$ the optimal investment of the arbitrageur solves

$$\max_{\{x\}} (\mathbb{E}[P_1|\hat{\theta}] - P_0)x - \frac{\rho}{2}^2 \mathbb{V}[P_1|\hat{\theta}].$$

After substituting for $\mathbb{E}[P_1|\hat{\theta}]$, $\mathbb{V}[P_1|\hat{\theta}]$ and P_0 from equations (33), (34) and (36) respectively, the arbitrageur's investment choice problem becomes

$$\max_{\{x\}} \left(\frac{m\tau_A(\hat{\theta}-\mu)}{m\tau_A + \tau_\theta} - \frac{\rho}{m\tau_A+\tau_\theta}x - \frac{\rho^3}{\tau_\theta^2}\sigma_u^2(x+u_0)\right) x - \frac{\rho}{2}x^2\sigma_P^2(m). \tag{37}$$

This objective function shows that the purchase x lowers the arbitrageur's expected payoff in three ways: via its negative impact on the expected resale price $\hat{P}_1$ at $t=1$, via its positive impact on its purchase price P_0, and via the increased risk borne by the arbitrageur. The first-order condition of the maximization problem (37) is

$$\frac{m\tau_A(\hat{\theta}-\mu)}{m\tau_A+\tau_\theta} - \frac{\rho^3}{\tau_\theta^2}\sigma_u^2 u_0 - \frac{\rho}{m\tau_A+\tau_\theta}2x - \frac{\rho^3}{\tau_\theta^2}\sigma_u^2 2x - \rho x \sigma_P^2(m) = 0.$$

Since assets are symmetric, this condition determines the arbitrageur's optimal investment in all advertised assets $x_i = x$, $i = 1,..,L$:

$$x = \frac{1}{\rho}\left(\frac{m\tau_A(\hat{\theta}-\mu)}{m\tau_A+\tau_\theta} - \frac{\rho^3}{\tau_\theta^2}\sigma_u^2 u_0\right)\frac{1}{2(\frac{1}{m\tau_A+\tau_\theta} + \frac{\rho^2}{\tau_\theta^2}\sigma_u^2) + \sigma_P^2(m)}. \tag{38}$$

Substituting for $\sigma_P^2(m) = \frac{\rho^2\sigma_u^2}{(m\tau_A+\tau_\theta)^2}$, the previous expression becomes

$$x = \frac{1}{\rho}\left(m\tau_A(\hat{\theta}-\mu) - (m\tau_A+\tau_\theta)\frac{\rho^3}{\tau_\theta^2}\sigma_u^2 u_0\right)\frac{1}{2(1+(m\tau_A+\tau_\theta)\frac{\rho^2}{\tau_\theta^2}\sigma_u^2) + \frac{\rho^2\sigma_u^2}{m\tau_A+\tau_\theta}}, \tag{39}$$

which can be used to write the arbitrageur's payoff from investing in an advertised asset:

$$\pi(m) = \frac{1}{4\rho}\left(m\tau_A(\hat{\theta}-\mu) - (m\tau_A+\tau_\theta)\frac{\rho^3}{\tau_\theta^2}\sigma_u^2 u_0\right)^2 \frac{1}{m\tau_A + \tau_\theta + \frac{\rho^2\sigma_u^2}{\tau_\theta^2}(m\tau_A+\tau_\theta)^2 + \frac{\rho^2\sigma_u^2}{2}}.$$

This expression is analogous to expression (20) for the arbitrageur's payoff obtained under the assumption of no price impact. As in that case, the arbitrageur benefits both from trading on his private information (the first term in parenthesis) and from exploiting the noise traders' orders at $t = 0$ (the second term). We again focus only on the contribution of advertising private information to the arbitrageur's profits, and thus neglect the second source of profits, i.e., set the noise trade shock $u_0 = 0$. Hence, the arbitrageur's expected payoff as a function of the fraction m of informed investors reduces to

$$\pi(m) = \frac{1}{4\rho} \frac{m^2 \tau_A^2 (\hat{\theta} - \mu)^2}{m\tau_A + \tau_\theta + \frac{\rho^2 \sigma_u^2}{\tau_\theta^2}(m\tau_A + \tau_\theta)^2 + \frac{\rho^2 \sigma_u^2}{2}}. \tag{40}$$

From expression (40), we obtain the arbitrageur's optimal advertising decision, noting that, upon advertising L symmetric assets, his total expected payoff is $L\pi(m(L))$, where $m(L) = \min[1, K/L]$. The following proposition characterizes the arbitrageur's optimal policy in terms of the fraction of investors that he chooses to inform with his ads:

Proposition 9. *If the arbitrageur's trades have price impact, the optimal fraction of informed investors, abstracting from integer constraints on K and L, is*

$$m^* = \min\left[\frac{\tau_\theta}{\tau_A}\sqrt{\frac{3}{2} + \frac{\tau_\theta}{\rho^2 \sigma_u^2}}, 1\right].$$

Hence, even when his trades have price impact, the arbitrageur may want to restrict his advertising activity exactly to the number of ads that an investor can process ($m^* = K/L = 1$), exactly as it is the case when his trades have no price impact. However, now there is also a parameter region where he will prefer to advertise a larger number of assets than under no price impact ($L > K$), thus lowering the probability that investors will process each ad below 1 ($m^* = K/L < 1$). This parameter region is defined by the following inequality:

$$\frac{\tau_\theta}{\tau_A}\sqrt{\frac{3}{2} + \frac{\tau_\theta}{\rho^2 \sigma_u^2}} < 1. \tag{41}$$

The intuitive reason behind the existence of this region is that, in the presence of price impact, the desire to mitigate price impact may induce the arbitrageur to advertise more than K assets, so as to spread his trades across a larger number of assets and thereby reduce the price impact for each of them. Hence, in this case the arbitrageur's objective of preventing dilution of investors' attention must be traded off against the gain from the increase in market liquidity that can be obtained by spreading trades across many assets. This intuition squares with the fact that the region defined by inequality (41) is larger when the risk aversion coefficient ρ and the variance of noise trading σ_u^2 increase: these parameter changes result in a stronger price impact of the arbitrageur's trades at $t = 0$ and $t = 1$, as can be seen from expressions (33) and (36). Indeed, if ρ and/or σ_u^2 are large, investors at $t = 0$ and at $t = 1$ are afraid of taking large positions in assets, which limits the liquidity of the market. Therefore, in these circumstances mitigating price takes precedence over avoiding dilution of investors' attention.

6 Conclusions

We conclude by summarizing the testable hypotheses about the investment and advertising activity of arbitrageurs that are generated by our model. Several of these predictions have already been shown to be consistent with some empirical evidence:

(i) Arbitrageurs concentrate advertising on a few assets at a time, depending on the available information processing capacity of investors. This is consistent with the fact that hedge fund managers that advertise their trading recommendations tend to target one company at a time, as documented by Ljungqvist and Qian (2016). Similarly, Della Corte et al. (2021) find that short-sellers whose trades are publicly disclosed have short positions that are concentrated on relatively few stocks, with more than 40% of them having a single disclosed net short position.

(ii) Advertising accelerates price discovery, and on average it increases arbitrageurs' profits: this prediction is consistent with the finding by Ljungqvist and Qian (2016) that on average the price of the stocks targeted by short-selling arbitrageurs in their sample drops by 7.5% on the date they release their first report, and by 21.4% to 26.2% in the three subsequent months, and with that by Luo (2018) that the stocks pitched by hedge funds at conferences outperform their benchmark by 7% in the subsequent 9 months, after earning a 20% cumulative abnormal return in the previous 18 months. Similarly, net short positions that are disclosed under European Short Selling Regulation (SSR 236/2012) are informative about subsequent asset price movements (Della Corte et al. (2021)).

(iii) Different arbitrageurs will tend to advertise the same opportunities and exploit them simultaneously, displaying a behavior sometimes referred to as "wolf pack". Zuckerman (2012) finds that, upon being publicly identified as overvalued by managers of large US equity hedge funds, stocks were shorted by several funds at once, either directly or via changes in put option exposures, and underperformed their benchmarks over the subsequent two years.

Other predictions of our model, instead, still await empirical testing:

(i) Arbitrageurs should overweight advertised assets in their portfolios, and such overweighting should be increasing in the precision of arbitrageurs' private information and the asset's expected appreciation, and may be decreasing in the variance of noise trading.

(ii) Arbitrageurs are more likely to advertise assets that are more severely mispriced, those for which their private information is more precise, and those whose information is easier to process for investors.

(iii) The more numerous are the arbitrageurs who simultaneously advertise different assets, the weaker is the price correction induced by each ad, and the lower is the risk-adjusted expected profit of each arbitrageur, because each ad will be processed by a smaller fraction of investors.

(iv) The prediction that arbitrageurs concentrate advertising on a few assets at a time is attenuated for large investors whose trades have significant price impact, as these may prefer to advertise their positions in several assets at the same time in order to mitigate their price impact.

Appendix

Proof of Proposition 2. If the number of advertised assets rises from L to $L+1$, then the arbitrageur's expected payoff (21) changes by

$$\Delta\Pi(L) = \frac{\tau_A^2(\hat{\theta}-\mu)^2}{2\rho^3\sigma_u^2}\left((L+1)\left(m(L+1)\right)^2 - L\left(m(L)\right)^2\right).$$

Increasing the number of advertised assets has a different impact on the arbitrageur's utility depending on whether investors' information capacity is saturated ($L = K$) or not ($L < K$). If it is not, then using expression (1) for $m(L)$, the arbitrageur's utility rises by

$$\Delta\Pi(L) = \frac{\tau_A^2(\hat{\theta}-\mu)^2}{2\rho^3\sigma_u^2}. \tag{42}$$

If instead investors' attention is already saturated, i.e. $L \geq K$, increasing the number of advertised assets from L to $L+1$ leads to a drop in the arbitrageur's utility:

$$\Delta\Pi(L) = -\frac{\tau_A^2(\hat{\theta}-\mu)^2}{2\rho^3\sigma_u^2}\frac{K^2}{(L+1)L}.$$

Hence, if the arbitrageur has information about $M > K$ assets, he will entirely use up investors' attention, but not over-exploit it. Obviously, if $M \leq K$, he will advertise all M assets he is informed about because $\Delta\Pi(L)$ is given by (42), which is positive for any $L \leq M$.

QED.

Proof of Proposition 3. We want to show that the arbitrageur chooses to advertise L^* assets such that $\sum_{i\in\mathbf{L}^*} D_i \leq C$. Suppose that, upon applying the algorithm described in the text before the proposition, each investor can process L but not $L+1$ messages, i.e., $\sum_{i=1}^{L} D_i \leq C$, but $\sum_{i=1}^{L+1} D_i > C$. If the arbitrageur advertises $L+1$ messages, then investors will randomly process only L of these $L+1$ messages. Hence, the fraction of informed investors per advertised asset will be $m = L/(L+1)$, so that the arbitrageur's expected payoff will be

$$\Pi(L+1) = \sum_{i=1}^{L+1} \pi_{Ai} = \sum_{i=1}^{L+1} k_i \tau_{Ai}^2 \frac{L^2}{(L+1)^2}, \tag{43}$$

where for convenience we use the shorthand

$$k_i \equiv \frac{(\hat{\theta}_i - \mu_i)^2}{2\rho^3 \sigma_{u_i}^2}.$$

As advertised assets contribute differently to the expected payoff of the arbitrageur, let us denote by j the advertised asset that contributes the least to his expected payoff:

$$j = \underset{i=1,..,L+1}{\arg\min} k_i \tau_{Ai}^2 \frac{2L+1}{(L+1)^2}. \tag{44}$$

Now suppose that the arbitrageur decides to drop asset j from the set of advertised assets, reducing their number to L and thus enabling investors to process all the messages, so that $m = 1$ for all advertised assets. The resulting expected payoff will be

$$\Pi(L) = \sum_{i=1}^{L} k_i \tau_{Ai}^2. \tag{45}$$

Let us denote by $\Delta\Pi \equiv \Pi(L) - \Pi(L+1)$ the change in the arbitrageur's expected payoff from dropping asset j from the set of $L+1$ advertised assets:

$$\Delta\Pi = \sum_{i=1}^{L} k_i \tau_{Ai}^2 \frac{2L+1}{(L+1)^2} - k_j \tau_{Aj}^2 \frac{L^2}{(L+1)^2}. \tag{46}$$

We need to show that $\Delta\Pi \geq 0$. By the definition of j in (44), each of the terms of the sum in expression (46) is weakly larger than an analogous term for asset j, so that

$$\sum_{i=1}^{L} k_i \tau_{Ai}^2 \frac{2L+1}{(L+1)^2} \geq L k_j \tau_{Aj}^2 \frac{2L+1}{(L+1)^2},$$

which implies that $\Delta\Pi \geq 0$ in (46) since

$$L k_j \tau_{Aj}^2 \frac{2L+1}{(L+1)^2} - k_j \tau_{Aj}^2 \frac{L^2}{(L+1)^2} = k_j \tau_{Aj}^2 \frac{L}{L+1} > 0.$$

This proves that the arbitrageur prefers to advertise L rather than $L+1$ assets. A similar argument can be used to show that the arbitrageur does not want to advertise more that K assets.

QED.

Proof of Proposition 7. First, consider a candidate equilibrium where all arbitrageurs advertise the same L bad assets, for $L \geq K$, so that investors' attention capacity is already saturated by information about them (as required by Proposition 2). For this to be an equilibrium, no arbitrageur must have the incentive to deviate from it by advertising a good asset.

If arbitrageur i follows an equilibrium strategy, his expected payoff from advertising the L bad assets is

$$\Pi_i = L\frac{K^2}{L^2}\pi_B.$$

If instead he deviates by advertising a good asset while other arbitrageurs keep advertising the L bad assets (so that advertised assets become $L+1$ in total), then his expected payoff becomes

$$\Pi_i' = L\frac{K^2}{(L+1)^2}\pi_B + \frac{K^2}{(L+1)^2}\pi_G.$$

Thus the arbitrageur will not deviate from the candidate equilibrium with L bad asset being advertised if

$$\Pi_i' - \Pi_i = \frac{K^2}{(L+1)^2}\left(\pi_G + \pi_B\frac{(L+1)^2}{L} - L\pi_B\right) \leq 0,$$

which is equivalent to

$$\pi_G \leq \pi_B\left(2 + \frac{1}{L}\right). \tag{47}$$

This shows that there is an equilibrium in which only bad assets are advertised if (47) holds, because no arbitrageur would prefer to deviate and advertise a good asset. Note that condition (30) implies (47) for any $L \leq AM$. Naturally, if arbitrageurs advertise any combination of good and bad assets in equilibrium, they get a higher expected payoff than from advertising bad assets only, which further decreases the appeal of deviating by advertising a different asset. Hence, any combination of $L \geq K$ assets can be advertised in equilibrium.

QED.

Proof of Proposition 8. It is easy to see that $\Pi(L_i; L_{-i})$ given by (31) increases with L_i if $L_{-i} + L_i \leq K$ because $m = 1$:

$$\Pi(L_i, L_{-i}) = L_i \frac{\tau_A^2(\hat{\theta} - \mu)^2}{2\rho^3\sigma_u^2},$$

so that in equilibrium it cannot be $L_{-i} + L_i < K$, as in this case some arbitrageur i would deviate by increasing L_i.

Next, notice that if $L_{-i} + L_i > K$, the derivative of $\Pi(L_i, L_{-i})$ with respect to L_i is

$$\frac{\partial \Pi(L_i; L_{-i})}{\partial L_i} = (L_{-i} - L_i)\frac{K^2}{(L_i + L_{-i})^3}\frac{\tau_A^2(\hat{\theta} - \mu)^2}{2\rho^3\sigma_u^2}.$$

In a symmetric equilibrium, $L_{-i} = (A-1)L^*$ and $L_i = L^*$, which for $A > 2$ implies $L_{-i} > L_i$ and $\frac{\partial \Pi(L_i; L_{-i})}{\partial L_i} \geq 0$. Hence, in equilibrium the arbitrageur would benefit from advertising additional assets, so that he advertises all M assets he is informed about.

QED.

Proof of Proposition 9. The arbitrageur acquires positions x_i, $i = 1, .., L$, in L advertised assets at $t = 0$, and at $t = 1$ liquidates them, i.e. trades $-x_i$. In deriving these initial positions, the advertising decision is taken as given. Later on we characterize the optimal advertising decision.

Building on expression (40), one can analyze the arbitrageur's optimal advertising decision. If the arbitrageur advertises $L \leq K$, then the fraction of informed investors about each advertised asset is $m = 1$, and his expected payoff is

$$\Pi = L\pi(1) = \frac{L}{4\rho}\tau_A^2(\hat{\theta} - \mu)^2 \frac{1}{(\tau_A + \tau_\theta) + (\tau_A + \tau_\theta)^2 \frac{\rho^2}{\tau_\theta^2}\sigma_u^2) + \frac{\rho^2\sigma_u^2}{2}}.$$

Since for $L \leq K$ the payoff increases linearly in L, at least K assets will be advertised. In principle, the arbitrageur may choose to advertise more assets. When $L \geq K$ assets are advertised, then the fraction of investors informed about each asset is $m = K/L \leq 1$, and the arbitrageur's payoff becomes

$$\Pi(L)=\frac{K}{4\rho}m\tau_A^2(\hat{\theta}-\mu)^2\frac{1}{m\tau_A+\tau_\theta+\frac{\rho^2\sigma_u^2}{\tau_\theta^2}(m\tau_A+\tau_\theta)^2+\frac{\rho^2\sigma_u^2}{2}},$$

which can be further simplified to

$$\Pi(L)=\frac{K}{4\rho}\tau_A^2(\hat{\theta}-\mu)^2\frac{1}{\tau_A+\frac{\tau_\theta}{m}+\frac{\rho^2\sigma_u^2}{\tau_\theta^2}(m\tau_A^2+2\tau_A\tau_\theta+\frac{\tau_\theta^2}{m})+\frac{\rho^2\sigma_u^2}{2m}}. \tag{48}$$

Denote the function in the denominator by $Z(m)=\tau_A+\frac{\tau_\theta}{m}+\frac{\rho^2\sigma_u^2}{\tau_\theta^2}(m\tau_A^2+2\tau_A\tau_\theta+\frac{\tau_\theta^2}{m})+\frac{\rho^2\sigma_u^2}{2m}$, and take its first and second derivatives:

$$Z'(m)=-\frac{\tau_\theta}{m^2}+\frac{\rho^2\sigma_u^2}{\tau_\theta^2}(\tau_A^2-\frac{\tau_\theta^2}{m^2})-\frac{\rho^2\sigma_u^2}{2m^2},$$

$$Z''(m)=2\frac{\tau_\theta}{m^3}+3\frac{\rho^2\sigma_u^2}{\tau_\theta^2}\frac{\tau_\theta^2}{m^3}+\frac{\rho^2\sigma_u^2}{m^3}>0.$$

The function $Z(m)$ is convex for $m\in(0,1]$, and therefore it reaches its minimum either at $m^*=\hat{m}$ such that $Z'(\hat{m})=0$, or at $m^*=1$ if $\hat{m}>1$. One can express the profit-maximizing fraction of informed investors, m^*, as follows:

$$m^*=\min\left[\frac{\tau_\theta}{\tau_A}\sqrt{\frac{3}{2}+\frac{\tau_\theta}{\rho^2\sigma_u^2}},1\right].$$

The minimum of the function $Z(m)$ corresponds to the arbitrageur's maximal expected payoff, so that he optimally chooses to advertise the number of assets L such that $\frac{K}{L}\approx m^*$. Specifically, he advertises $L=K$ if $m^*=\frac{\tau_\theta}{\tau_A}\sqrt{\frac{3}{2}+\frac{\tau_\theta}{\rho^2\sigma_u^2}}\geq 1$. If $m^*<1$, he chooses between $\underline{L}=\lfloor\frac{K}{m^*}\rfloor$ and $\overline{L}=\lceil\frac{K}{m^*}\rceil$. Formally, he advertises $\underline{L}$ assets if

$$\tau_\theta\frac{\overline{L}}{K}+\frac{\rho^2\sigma_u^2\tau_A^2}{\tau_\theta^2}\frac{K}{\overline{L}}+\frac{3}{2}\rho^2\sigma_u^2\frac{\overline{L}}{K}\geq\tau_\theta\frac{\underline{L}}{K}+\frac{\rho^2\sigma_u^2\tau_A^2}{\tau_\theta^2}\frac{K}{\underline{L}}+\frac{3}{2}\rho^2\sigma_u^2\frac{\underline{L}}{K},$$

and he advertises $\overline{L}$ otherwise.

QED.

Online Appendix

Attention Allocation and Learning Accuracy

As explained in the text of the paper, Proposition 2 (and more specifically Corollary 2) by Van Nieuwerburgh and Veldkamp (2010) implies that in our setting each investor will optimally invest his entire attention capacity in processing K messages with perfect accuracy rather than attempting to process more than K messages with imperfect accuracy. To illustrate this point, in what follows we focus on an example with two symmetric assets, where the investor is capable of perfectly processing advertised information about a single asset, i.e., $K = 1$, or imperfectly learning about both assets, i.e., $L = 2$. We show that the investor will optimally choose to process with perfect accuracy the ad regarding a single asset (randomly chosen as the two assets are ex-ante symmetric), rather than learning imperfectly about both.

Since the two assets are assumed to be symmetric, in what follows subscripts to distinguish them will be used only where necessary. Recall that in our setting investors' prior about the future payoff of an asset is μ, and has precision $\tau_\theta = 1/\sigma_\theta^2$. The precision of the arbitrageur's signal is τ_A. If an investor perfectly learns the arbitrageur's signal, the conditional distribution of the asset's future payoff is $N(\theta_I, \sigma_I^2)$, where $\theta_I = \mathbb{E}[\theta|\hat{\theta}] = \frac{\tau_A\hat{\theta}+\tau_\theta\mu}{\tau_A+\tau_\theta}$, and $\sigma_I^2 = \mathbb{V}[\theta|\hat{\theta}] = \frac{1}{\tau_A+\tau_\theta}$.

The total variance of the future payoff for investors who perfectly learn about an advertised asset is $\sigma_I^2 = \frac{1}{\tau_A+\tau_\theta}$, while for uninformed investors it is $\sigma_U^2 = \frac{1}{\tau_\theta}$, so that advertising reduces the variance by

$$\Delta \equiv \sigma_U^2 - \sigma_I^2 = \frac{1}{\tau_\theta} - \frac{1}{\tau_A+\tau_\theta} = \frac{\tau_A}{\tau_\theta(\tau_A+\tau_\theta)}. \tag{49}$$

Hence, Δ measures the extent to which an investor can reduce the uncertainty about the future payoff of an asset by allocating his entire attention capacity to information regarding that asset. As such, it is equivalent to the definition of attention capacity in Van Nieuwerburgh and Veldkamp (2010) as the maximum entropy reduction attainable by an investor. If instead an investor decides to partially process the ads issued about two assets, i.e. learn about each of them with precision $\hat{\tau}$, then the total reduction in the payoff uncertainty attainable by fully using his attention capacity will be

$$\frac{2}{\tau_\theta} - \frac{2}{\hat{\tau}} \leq \Delta.$$

When the above condition holds with equality, one can use it jointly with (49) to determine the maximal precision $\hat{\tau}$ with which investors can imperfectly learn signals when they split their learning capacity among the two assets:

$$\hat{\tau} = \frac{2\tau_\theta}{2 - \frac{\tau_A}{\tau_A + \tau_\theta}}.$$

Since the resulting investor's information about the two advertised assets combines the investor's own prior and the arbitrageur's signals with some noise, we can alternatively write the precision of the investor's posterior as

$$\hat{\tau} = \tau_\theta + \hat{\tau}_A,$$

where $\hat{\tau}_A$ denotes the precision with which the investor learns each of the two signals advertised by the arbitrageur. Combining the two previous expressions yields the maximal value of this precision attainable by using entirely one's information capacity:

$$\hat{\tau}_A = \hat{\tau} - \tau_\theta = \frac{\frac{\tau_A \tau_\theta}{\tau_A + \tau_\theta}}{2 - \frac{\tau_A}{\tau_A + \tau_\theta}} = \frac{\tau_\theta}{2\tau_\theta + \tau_A} \tau_A < \frac{\tau_A}{2}. \tag{50}$$

Thus, if the investor attempts to partially process advertising about two assets, then he gets two noisy signals with low precision $\hat{\tau}_A$ rather than a single signal with a high precision τ_A. In what follows we will denote this noisy signal by $\theta' = \hat{\theta} + \nu = \theta + \varepsilon + \nu$, so that $\nu \sim N(0, 1/\hat{\tau}_A - 1/\tau_A)$.

Now we compare the investor's expected profits in two situations: (i) when he perfectly learns the arbitrageur's signal about a single asset and relies on the prior for the other asset, and (ii) when he imperfectly learns the signals issued by the arbitrageur about both assets. In drawing this comparison, we assume that other investors are pursuing the attention allocation strategy (i). We first compute the investor's expected profits from pursuing strategy (i), and then verify that he is not better off by deviating to strategy (ii).

(i) If each investor learns about a single asset, then at $t = 1$ he solves the following portfolio choice problem, where the asset about which he chooses to learn is indexed by I and the other asset is indexed by U:

$$\max_{\{y_I, y_U\}} (\mathbb{E}(\theta|\hat{\theta}) - P_{1I})y_I + (\mathbb{E}(\theta) - P_{1U})y_U - \frac{\rho}{2}(\mathbb{V}(\theta|\hat{\theta})y_I^2 + \mathbb{V}(\theta)y_U^2), \tag{51}$$

Assume, with no generality, that the investor decides to learn about asset 1. Then, his optimal investments in the two assets are the same as in expressions (4) and (6):

$$y_I = (\tau_A\hat{\theta} + \tau_\theta\mu - (\tau_A + \tau_\theta)P_{1I})/\rho, \tag{52}$$

$$y_U = \tau_\theta(\mu - P_{1U})/\rho. \tag{53}$$

Note that if two assets are advertised and all investors (except possibly the deviating one) randomly pick one signal to process, then the fraction of informed investors for each advertised asset is $m = 1/2$. Recall that for $m = 1/2$ the equilibrium price for each advertised asset is given by (7).

Thus the equilibrium profit obtained at $t = 2$ from asset I about which the investor chooses to learn is

$$\begin{aligned} (\theta - P_{1I})y_I &= \left(\theta - \frac{\tau_A\hat{\theta}/2 + \tau_\theta\mu + \rho u}{\tau_A/2 + \tau_\theta}\right)\left(\tau_A\hat{\theta} + \tau_\theta\mu - (\tau_A + \tau_\theta)\frac{\tau_A\hat{\theta}/2 + \tau_\theta\mu + \rho u}{\tau_A/2 + \tau_\theta}\right)\frac{1}{\rho} \\ &= \frac{(-\tau_A\varepsilon/2 + \tau_\theta\eta - \rho u)(\tau_\theta\tau_A(\eta + \varepsilon)/2 - (\tau_A + \tau_\theta)\rho u)}{(\tau_A/2 + \tau_\theta)^2\rho}, \end{aligned} \tag{54}$$

where $\varepsilon = \hat{\theta} - \theta$, $\eta = \theta - \mu$, and ε, η and u are all independent. Taking the expected value of this expression and using the fact that $\tau_\theta = 1/\sigma_\theta^2$ and $\tau_A = 1/\sigma_\varepsilon^2$, one can express the investors' expected profit from investing in asset I as follows:

$$\frac{-\tau_A^2\tau_\theta\sigma_\varepsilon^2/4 + \tau_\theta^2\tau_A\sigma_\theta^2 + (\tau_A + \tau_\theta)\rho^2\sigma_u^2}{(\tau_A/2 + \tau_\theta)^2\rho} = \frac{\tau_\theta\tau_A 3/4 + (\tau_A + \tau_\theta)\rho^2\sigma_u^2}{(\tau_A/2 + \tau_\theta)^2\rho}.$$

The investor's profit at $t = 2$ from the other advertised asset U, about which he does not process information, is

$$(\theta - P_{1U})y_U = \left(\theta - \frac{\tau_A\hat{\theta}/2 + \tau_\theta\mu + \rho u}{\tau_A/2 + \tau_\theta}\right)\left(\mu - \frac{\tau_A\hat{\theta}/2 + \tau_\theta\mu + \rho u}{\tau_A/2 + \tau_\theta}\right)\frac{\tau_\theta}{\rho}$$
$$= \frac{(-\tau_A\varepsilon/2 + \tau_\theta\eta - \rho u)\tau_\theta(\tau_A(-\eta - \varepsilon)/2 - \rho u)}{(\tau_A/2 + \tau_\theta)^2\rho}. \tag{55}$$

The expected value of expression (55) is

$$\tau_\theta\frac{\tau_A^2\sigma_\varepsilon^2/4 - \tau_A\tau_\theta\sigma_\theta^2/2 + \rho^2\sigma_u^2}{(\tau_A/2 + \tau_\theta)^2\rho} = \tau_\theta\frac{-\tau_A/4 + \rho^2\sigma_u^2}{(\tau_A/2 + \tau_\theta)^2\rho}.$$

Hence, the investor's total expected profit from this strategy is

$$\frac{\tau_\theta\tau_A 3/4 + (\tau_A + \tau_\theta)\rho^2\sigma_u^2}{(\tau_A/2 + \tau_\theta)^2\rho} + \tau_\theta\frac{-\tau_A/4 + \rho^2\sigma_u^2}{(\tau_A/2 + \tau_\theta)^2\rho} = \frac{\tau_\theta\tau_A/2 + \rho\sigma_u^2(\tau_A + 2\tau_\theta)}{(\tau_A/2 + \tau_\theta)^2}. \tag{56}$$

(ii) Now consider the portfolio choice problem faced by an investor who decides to deviate and learn imperfectly about both assets being advertised. Such an investor will condition on two signals θ' of precision $\hat{\tau}_A < \tau_A$, and his optimization problem will be

$$\max_{\{y_1,y_2\}} \ (\mathbb{E}(\theta|\theta_1') - P_{11})y_1 + (\mathbb{E}(\theta|\theta_2') - P_{12})y_2 - \frac{\rho}{2}(\mathbb{V}(\theta|\theta_1')y_1^2 + \mathbb{V}(\theta|\theta_2')y_2^2). \tag{57}$$

Since assets are symmetric, we consider the optimal investment in one of them, denoting it by y_I', without indexing it. So the optimal investment in the asset is:

$$y_I' = (\hat{\tau}_A\theta' + \tau_\theta\mu - (\hat{\tau}_A + \tau_\theta)P_1)/\rho, \tag{58}$$

Thus the investor's final profit at $t = 2$ from investing in a single asset is

$$(\theta - P_1)y_I' = (\theta - \frac{\tau_A\hat{\theta}/2 + \tau_\theta\mu + \rho u}{\tau_A/2 + \tau_\theta})\left(\hat{\tau}_A\theta' + \tau_\theta\mu - (\hat{\tau}_A + \tau_\theta)\frac{\tau_A\hat{\theta}/2 + \tau_\theta\mu + \rho u}{\tau_A/2 + \tau_\theta}\right)\frac{1}{\rho},$$

which can be rewritten as

$$\frac{(-\tau_A\varepsilon/2 + \tau_\theta\eta - \rho u)(\hat{\tau}_A(\tau_A/2 + \tau_\theta)\nu - \tau_\theta(\hat{\tau}_A - \tau_A/2)\varepsilon + \tau_\theta(\hat{\tau}_A - \tau_A/2)\eta - (\hat{\tau}_A + \tau_\theta)\rho u)}{(\tau_A/2 + \tau_\theta)^2\rho}.$$

Taking the expectation of this expression yields the investor's expected profit:

$$\frac{\tau_A\tau_\theta(\hat{\tau}_A-\tau_A/2)\sigma_\varepsilon^2/2+\tau_\theta^2(\hat{\tau}_A-\tau_A/2)\sigma_\theta^2+(\hat{\tau}_A+\tau_\theta)\rho^2\sigma_u^2}{(\tau_A/2+\tau_\theta)^2\rho}=\frac{3\tau_\theta(\hat{\tau}_A-\tau_A/2)/2+(\hat{\tau}_A+\tau_\theta)\rho^2\sigma_u^2}{(\tau_A/2+\tau_\theta)^2\rho},$$

so that the total expected profit from investing in the two assets is

$$\frac{3\tau_\theta(\hat{\tau}_A-\tau_A/2)+2(\hat{\tau}_A+\tau_\theta)\rho^2\sigma_u^2}{(\tau_A/2+\tau_\theta)^2\rho}. \tag{59}$$

Subtracting expression (56) from expression (59) yields the change in expected profit from deviating from strategy (i) to strategy (ii):

$$\frac{(\hat{\tau}_A-\tau_A/2)(3\tau_\theta+2\rho^2\sigma_u^2)-\tau_\theta\tau_A/2}{(\tau_A/2+\tau_\theta)^2\rho}, \tag{60}$$

which is negative because $\hat{\tau}_A<\tau_A/2$ according to inequality (50).

This proves that the investor will prefer to use his entire information processing capacity to perfectly learn about a single asset rather than imperfectly about two. Hence, the logic behind Proposition 2 of Van Nieuwerburgh and Veldkamp (2010) extends to our setup. This proposition implies that, if investors have mean-variance preferences and assets have independent returns, it is not optimal for an investor to split his attention among several assets: he will concentrate his learning activity on a single asset. The same holds true in our setting: as illustrated by this example, it is not optimal for an investor to split his attention between two assets and get two noisy signals. This shows that our assumption that investors do not split their attention across assets entails no loss of generality: even in a setting where they could split their attention capacity across assets, they are better off concentrating it so as to perfectly learn the arbitrageur's signal about the corresponding asset.

References

Abreu, D. and M. Brunnermeier (2002): "Synchronization risk and delayed arbitrage," *Journal of Financial Economics*, 66, 341–360.

Barber, B. M. and T. Odean (2008): "All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors," *The Review of Financial Studies*, 21, 785–818.

Becht, M., J. Franks, J. Grant, and H. F. Wagner (2017): "Returns to Hedge Fund Activism: An International Study," *The Review of Financial Studies*, 30, 2933–2971.

Benabou, R. and G. Laroque (1992): "Using Privileged Information to Manipulate Markets: Insiders, Gurus, and Credibility," *The Quarterly Journal of Economics*, 107, 921–958.

Bikhchandani, S., D. Hirshleifer, and I. Welch (1992): "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades," *Journal of Political Economy*, 100, 992–1026.

Bommel, J. V. (2003): "Rumors," *The Journal of Finance*, 58, 1499–1519.

Brav, A., A. Dasgupta, and R. Mathews (2021): "Wolf Pack Activism," *Management Science*, published online.

Brunnermeier, M. K. (2005): "Information Leakage and Market Efficiency," *The Review of Financial Studies*, 18, 417–457.

Chen, H., P. De, Y. Hu, and B.-H. Hwang (2014): "Wisdom of Crowds: The Value of Stock Opinions Transmitted Through Social Media," *Review of Financial Studies*, 27, 1367–1403.

Della Corte, P., R. Kosowski, and N. Rapanos (2021): "Best Short," *SSRN Working Paper No. 3436433.*

DellaVigna, S. and J. M. Pollet (2009): "Investor Inattention and Friday Earnings Announcements," *The Journal of Finance*, 64, 709–749.

DeLong, J. B., A. Shleifer, L. H. Summers, and R. J. Waldmann (1990): "Noise Trader Risk in Financial Markets," *Journal of Political Economy*, 98, 703–738.

Froot, K. A., D. S.Scharfstein, and J. C. Stein (1992): "Herd on the Street: Informational Inefficiencies in a Market with Short-Term Speculation," *The Journal of Finance*, 47, 1461–1484.

Gromb, D. and D. Vayanos (2010): "Limits of Arbitrage," *Annual Review of Financial Economics*, 2, 251–275.

Huberman, G. and T. Regev (2001): "Contagious Speculation and a Cure for Cancer: A Nonevent that Made Stock Prices Soar," *The Journal of Finance*, 56, 387–396.

Indjejikian, R., H. Lu, and L. Yang (2014): "Rational Information Leakage," *Management Science*, 60, 2762–2775.

Lipnowski, E., L. Mathevet, and D. Wei (2020): "Attention Management," *American Economic Review: Insights*, 2, 17–32.

Liu, Y. (2017): "Why Do Large Investors Disclose Their Information?" *Swiss Finance Institute Research Paper No. 18-17.*

Ljungqvist, A. and W. Qian (2016): "How Constraining Are Limits to Arbitrage?" *The Review of Financial Studies*, 29, 1975–2028.

Luo, P. (2018): "Talking Your Book: Evidence from Stock Pitches at Investment Conferences," *SSRN Working paper No. 3190837.*

Pasquariello, P. and Y. Wang (2021): "Speculation with Information Disclosure," *SSRN Working paper No. 2847321.*

Peng, L. and W. Xiong (2006): "Investor Attention, Overconfidence and Category Learning," *Journal of Financial Economics*, 80, 563–602.

Scharfstein, D. S. and J. C. Stein (1990): "Herd Behavior and Investment," *The American Economic Review*, 80, 465–479.

Schmidt, D. (2019): "Stock Market Rumors and Credibility," *The Review of Financial Studies*, 33, 3804–3853.

Shleifer, A. and R. W. Vishny (1997): "The Limits of Arbitrage," *The Journal of Finance*, 52, 35–55.

Van Nieuwerburgh, S. and L. Veldkamp (2009): "Information Immobility and the Home Bias Puzzle," *The Journal of Finance*, 64, 1187–1215.

——— (2010): "Information Acquisition and Under-Diversification," *Review of Economic Studies*, 77, 779–805.

Veldkamp, L. (2011): *Information choice in macroeconomics and finance*, Princeton University Press.

Zuckerman, R. (2012): "Synchronized Arbitrage and the Value of Public Announcements," *Working paper*.