ADVERTISING ARBITRAGE

Sergey Kovbasyuk
Marco Pagano

Working Paper No 277

NES Working Paper series

August 2020

Advertising Arbitrage*

Sergey Kovbasyuk † Marco Pagano ‡

July 16, 2020

Abstract

Arbitrageurs with a short investment horizon gain from accelerating price discovery by advertising their private information. However, advertising many assets may overload investors' attention, reducing the number of informed traders per asset and slowing price discovery. So arbitrageurs optimally concentrate advertising on just a few assets, which they overweight in their portfolios. Unlike classic insiders, advertisers prefer assets with the least noise trading. If several arbitrageurs share information about the same assets, inefficient equilibria can arise, where investors' attention is overloaded and substantial mispricing persists. When they do not share, the overloading of investors' attention is maximal.

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

JEL classification: G11, G14, G2, D84.

*We are grateful to 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. Similarly, Zuckerman (2012) documents excess returns for hedge funds that publicly announce their short sales, and 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, owing to the arrival of noisy information, 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 arbitrageurs typically go well beyond simply stating their recommendation: they produce hard evidence buttressing it during their pitches, and

[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.

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 (2018).

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 overcoming the limited attention of rational investors, it reduces the risk incurred by the arbitrageur in liquidating his position. This risk arises either from noise traders or from independent public news that pushes the price away from the arbitrageur's private information. This risk reduction in turn enhances the arbitrageur's willingness to make large bets on his private information, engendering a complementarity between advertising and portfolio bias towards the securities advertised. 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:[4] 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 c losing the position profitably. Indeed, in practice hedge fund managers who advertise their recommendations typically pitch a single asset.

Second, concentrated advertising produces portfolio under-diversification. By lowering liquidation risk, advertising a given mispriced asset raises risk-adjusted expected return to the arbitrageur, so a risk-averse arbitrageur will want to overweight that asset in his

[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), who provide a model of gurus, assume them to be honest with a given probability and opportunistic with the complementary probability. If the guru is opportunistic and gets positive private information about an asset, he sends a negative message that drives the price down, buys cheap and gets a high return. Benabou and Laroque conclude that such gurus can still manipulate markets if they have some reputational capital. In contrast, the advertising arbitrageurs appear to push prices closer to fundamental values.

[4]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."

portfolio more than he would solely on the basis of his private information.

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 and his private information is most precise. He will also prefer an asset with little noise trading – a prediction in contrast with that of the standard insider trading model: here noise trading is the source of liquidation risk, while for the classic insider it is the source of informational rents. Moreover, the arbitrageur will want to time his advertising, concentrating it when investors' attention capacity is greatest, i.e. when they are not distracted by major aggregate news, and so are most receptive to his pitch.

Fourthly, if arbitrageurs share information about the same set of mispriced assets, they tend to feature "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, they may get collectively trapped in a situation where investors attention is overloaded by information about too many assets, or 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.

Finally, if arbitrageurs do not share information, i.e. if they have exclusive private information about different assets, they always overload investors' attention to the extreme, so that collectively they have lower expected payoff and greater liquidation risk than under information sharing. As they have no incentive to coordinate, they end up over-exploiting a common scarce resource, i.e. 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.

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, investors' limited attention is the reason for advertising, which succeeds precisely when it catches the attention of investors, i.e. when it induces them to devote their limited processing capacity to the opportunity identified.[5] 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 hypothetically 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 even more strongly towards the advertised assets than their private information alone would warrant.

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 Froot et al. (1992) and Bikhchandani et al. (1992). But in our model herding arises from the strategic complementarity in advertising and investing by arbitrageurs, and speeds up price discovery. By contrast, in informational cascades investors disregard their own information in favor of inference based on 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, but 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, taking the decisions of the arbitrageur as given, and the resulting equilibrium prices of assets. Section 3 characterizes the arbitrageur's optimal advertising and investment decisions, studies how asset characteristics affect the gain from advertising them, compares the results with those of a classic insider trading model. Next, Section 4 considers multiple informed arbi-

[5] The same result would obtain if information about mispricing were costly: in this case advertising would work not by directing investors' attention to information but by conveying it free of charge. So our model can be reinterpreted as one where advertising facilitates costly information acquisition by investors.

trageurs, and allowing for strategic interactions among them. Finally, Section 5 relaxes an important assumption maintained 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.

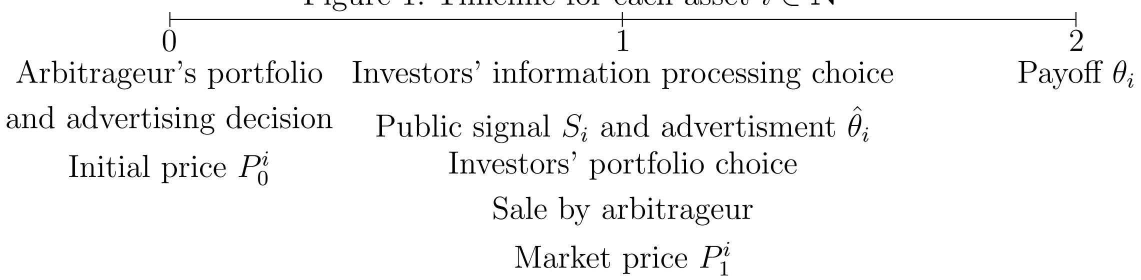
1 The Model

In the baseline model we consider an economy with a continuum of risky assets $\mathbf{N}$, available in zero net supply, and a safe asset that for simplicity is assumed to pay zero interest. 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 also relax the assumption of price-taking behavior by arbitrageurs.

Preferences. Rational investors have constant absolute risk-aversion (CARA) preferences: their utility from a monetary payoff c is $V(c) = 1 - e^{-\rho c/2}$. The parameter $\rho/2 > 0$ is the Arrow-Pratt measure of the absolute risk aversion. Noise traders trade each asset $i \in N$, and their total demand is $u_i \sim N(0, \sigma^2_{u_i})$.

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^i_0 and P^i_1 respectively, and delivers a final payoff θ_i at $t = 2$.

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 $\tau^A_i = 1/\sigma^2_{\epsilon_i}$, $i \in \mathbf{M}$. The arbitrageur has no private information about assets that do not belong to the set $\mathbf{M}$. If he takes a non-zero position x_i in asset i at $t = 0$, he cannot wait until its final payoff is realized at $t = 2$: he must liquidate his position at $t = 1$. This captures either the urgency of investing in other profitable assets or the short-termism of fund mangers due to high-powered incentive compensation.[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, except via advertising. From their point of view, any asset $i \in \mathbf{N}$ can be in $\mathbf{M}$ with the same probability, and since the set of traded assets is a continuum, this probability is zero. This implies that, unlike standard models of informed trading, this model posits no learning from prices: investors can learn about fundamentals only from public signals or from arbitrageurs' advertising.

Investors receive free public information at $t = 1$ about the future payoff of asset i in the form of a signal $S_i = \theta_i + \eta_i$, $i \in \mathbf{N}$: $\eta_i \sim N(0, \sigma^2_{\eta_i})$. The precision of the public signal is $\tau_i^S = 1/\sigma^2_{\eta_i}$. One can think of the public signal as information about a single market-wide factor, so that individual signals S_i, $i \in \mathbf{N}$, capture the effect of this market-wide factor on each asset i. We assume the errors η_i and ε_i to be independent for any $i \in \mathbf{M}$, so that the arbitrageur's informational advantage does not lie just in observing public information before other investors, as in the case of an insider, but in observing an independent signal about the future asset payoff, which complements public information. As we will see, this feature makes the arbitrageur's behavior very different from that of a classic insider.

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 therefore market prices. For simplicity, L denotes 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.[7] Hence, investors who do not process the signal advertised by the arbitrageur must rely solely on public signals in their portfolio choices.

[6] Alternatively, one can think of the arbitrageur as incurring holding costs, as in Abreu and Brunnermeier (2002), so that he prefers to liquidate without waiting for the final payoff.

[7] Note that, even if they did know, inferring the arbitrageur's signal from the relevant market price would be no less costly than processing the signal itself.

Limited attention. Each investor has limited attention. In the spirit of Van Nieuwerburgh and Veldkamp (2010), we assume that there is a limit to the amount of information an investor can process. We model this limited attention capacity as the maximum number K of signals about different assets that can be processed over and above public signals S_i, $i \in \mathbf{N}$. Insofar as processing public signals also requires attention, K can be considered as the residual attention capacity of investors, after processing public information.[8] If the arbitrageur advertises $L > K$ assets, each investor can learn only about K assets, taken randomly in $\mathbf{L}$.

In what follows, we refer to the mass m_i of investors who learn from the advertisement of asset i as "informed" and to the remaining $1 - m_i$ investors as "uninformed". For instance, if the arbitrageur advertises $L > K$ assets, and each investor learns about a random sample of K assets in $\mathbf{L}$, then 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 explain, the arbitrageur gains by disseminating his information to many informed investors m_i: 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. an asset that can be advertised by the arbitrageur. To simplify the notation, we drop the index i wherever this can be done with no loss of clarity. For instance, we shall refer to the mass of investors informed about asset i at $t = 1$ as $m \in [0, 1]$.

We focus on the determination of the prices of advertised assets at $t = 1$: prices for other assets can be obtained as a limiting special case where $m = 0$. At $t = 1$, the price P_1 must clear the market for the asset, balancing the net demand brought by the arbitrageur,

[8] Insofar as public information concerns a market-wide factor affecting the value of all assets, it is natural to assume that processing it takes priority over processing asset-specific information advertised by arbitrageurs.

noise traders, and informed and uninformed investors.

At $t = 1$ the investors' optimal strategy consists of an attention allocation decision and a 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 signals, 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 by investors and the resulting fraction $m \in (0, 1]$ of informed investors as given, we can study investors' portfolio choice and characterize the market clearing price for each advertised asset.

Informed investors condition their demand y_I on both the public signal S and the arbitrageur's signal $\hat{\theta}$, so that from their point of view at $t = 1$ the conditional distribution of the asset's future payoff is $N(\theta_I, \sigma_I^2)$, where $\theta_I = \mathbb{E}[\theta|S, \hat{\theta}] = \frac{\tau_A\hat{\theta}+\tau_S S}{\tau_A+\tau_S}$, and $\sigma_I^2 = \text{var}[\theta|S, \hat{\theta}] = \frac{1}{\tau_A+\tau_S}$. Since investor's utility is CARA and asset payoffs are normal, each of them will maximize the certainty equivalent of their payoff:

$$\max_{\{y_I\}} \; (\theta_I - P_1)y_I - \frac{\rho}{2}y_I^2\sigma_I^2. \tag{2}$$

Hence, each of them buys

$$y_I = (\tau_A\hat{\theta} + \tau_s S - (\tau_A + \tau_S)P_1)/\rho. \tag{3}$$

Uninformed investors solve a problem similar to that of the informed, with the key difference that they condition their demand solely on the public signal S:

$$\max_{\{y_U\}} \; (\theta_U - P_1)y_U - \frac{\rho}{2}y_U^2\sigma_U^2. \tag{4}$$

From their point of view, at $t = 1$ the conditional distribution of the asset's future payoff is $N(\theta_U, \sigma_U^2)$, where $\theta_U = \mathbb{E}[\theta|S] = S$, and $\sigma_U^2 = \text{var}[\theta|S] = 1/\tau_S$. Each uninformed investor buys

$$y_U = \tau_s(S - P_1)/\rho. \tag{5}$$

As there are m informed investors and $1 - m$ uninformed ones, noise traders buy an amount u, and the arbitrageur's trade is negligible, market clearing requires

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

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

$$P_1 = \frac{m\tau_A\hat{\theta} + \tau_S S}{m\tau_A + \tau_S} + \frac{\rho}{\tau_S + m\tau_A}u. \tag{6}$$

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 (low 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 (6) whatever initial position he may have acquired at $t = 0$.[9] By the same token, the arbitrageur considers the price at $t = 0$, P_0, as given. For assets that are not advertised, the equilibrium price at $t = 1$ is obtained by setting $m = 0$ in (6), i.e. $P_1 = S + \frac{\rho}{\tau_S}u$.

3 The arbitrageur's decisions

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 $t = 1$ conditional on his portfolio. Finally, we characterize his optimal portfolio under optimal advertising.

3.1 Investment decision

As of $t = 0$, the arbitrageur has no private information about assets outside the set $\mathbf{M}$. He might choose to trade these assets or not depending on their prices and public information; in either case these trades are of no interest because he has no information to advertise and behaves like a standard investor. In what follows we focus only on assets that he is privately informed about. Suppose the arbitrageur trades an amount x of an asset for which he has a private signal $\hat{\theta}$. For the time being, we take the advertising decision as given, and hence also the mass m of investors who process the advertisement about the

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

asset. As of $t = 0$, the arbitrageur expects the price P_1 from (6) to be distributed according to $P_1 \sim N(\hat{P}, \sigma_P^2(m))$, where

$$\hat{P} = \mathbb{E}[P_1|\hat{\theta}] = \hat{\theta},$$

and

$$\sigma_P^2(m) = \mathrm{var}[P_1|\hat{\theta}] = \frac{\tau_S + \tau_S^2/\tau_A + \sigma_u^2\rho^2}{(\tau_S + m\tau_A)^2}. \tag{7}$$

The arbitrageur will liquidate his position x in any asset at $t = 1$. Hence, his returns from investing in different assets are normal and independently distributed, so that at $t = 0$ his investment problem is equivalent to choosing his position x in any asset so as to maximize the resulting certainty-equivalent profit:

$$\max_{\{x\}} \; (\hat{P} - P_0)x - \frac{\rho}{2}x^2\sigma_P^2(m). \tag{8}$$

Hence his optimal investment is

$$x = \frac{\hat{\theta} - P_0}{\rho\sigma_P^2(m)} = \frac{\hat{\theta} - P_0}{\rho}\frac{(\tau_S + m\tau_A)^2}{\tau_S + \tau_S^2/\tau_A + \sigma_u^2\rho^2}. \tag{9}$$

Substituting for x in (8) yields the arbitrageur's expected gain from investing in an asset:

$$u(m) = \frac{(\hat{\theta} - P_0)^2}{2\rho}\frac{(\tau_S + m\tau_A)^2}{\tau_S + \tau_S^2/\tau_A + \sigma_u^2\rho^2}. \tag{10}$$

This expected gain is increasing in the fraction of investors m informed about his signal. This highlights the complementarity between advertising and investment, which as we shall see is a key feature of this model. Clearly, the arbitrageur benefits from the initial mispricing $\hat{\theta} - P_0$ and from the precision of his information τ_A, while he is harmed by variance of the noise trading σ_u^2. This is precisely the opposite of a classic insider, who benefits from the activity of noise traders. The reason is that noise traders decrease the impact of advertising on the price and increase the risk of the arbitrageur's position.

3.2 Optimal advertising

From now on, the assets that can be advertised are indexed by $i \in \mathbf{M}$, but are assumed to be symmetric, in the sense that their returns are identically and independently distributed (the case of heterogeneous assets is analyzed in Section 3.3). If the arbitrageur advertises L assets, then from (1) the mass of investors that pay attention to his advertising about each asset is $m(L) = \min[1, K/L]$. To simplify notation, we introduce the following shorthand:

$$k \equiv \frac{(\hat{\theta} - P_0)^2}{2\rho} \frac{1}{\tau_S + \frac{\tau_S{}^2}{\tau_A} + \sigma_u^2 \rho^2}, \tag{11}$$

so that expression (10), i.e. the arbitrageur's expected payoff from an optimal position in any advertised asset $i \in \mathbf{L}$, can be rewritten as

$$u_i^A = k(\tau_S + \tau_A m(L))^2, \tag{12}$$

and his expected payoff from an optimal position in a non-advertised asset $j \in \mathbf{M} \setminus \mathbf{L}$ as

$$u_j^N = k\tau_S{}^2. \tag{13}$$

The arbitrageur chooses the number of assets advertised L (out of the set $\mathbf{M}$ he is informed about) so as to maximize his expected payoff

$$U(L) = \sum_{i \in L} u_i^A + \sum_{j \in M \setminus L} u_j^N = kL\left[(\tau_S + \tau_A m(L))^2 - \tau_S^2\right] + M\tau_S^2, \tag{14}$$

which yields the following result:

Proposition 1 (**Concentrated advertising**). *The arbitrageur advertises* $\min(K, M)$ *assets, so that all investors are informed about them* ($m = 1$).

The reason why the arbitrageur will advertise no fewer than K assets is intuitive: leaving investors' attention capacity unexploited would be wasteful, since it would imply that the arbitrageur's information would not be impounded in market prices for some assets. As a result, the arbitrageur's profits from trading these assets would be riskier and lower than feasible, on average.

What is less intuitive is why the arbitrageur does not want to advertise more than K assets: after all, one might think that investors can always disregard the information that they cannot digest. However, in this model the arbitrageur has the incentive not to over-exploit investor's attention. To understand why, note that if the arbitrageur advertised more than K assets, each ad would compete for investors' attention against the others. As attention is split evenly among advertised assets, this would lower the mass of informed traders for each, and so moderate the price correction induced by each ad. Conversely, the more investors pay attention to the advertisement of an asset, the closer its price will be to the arbitrageur's estimate of its fundamental value at $t = 1$, and therefore the less risky the arbitrageur's position in the asset, and the greater his gain from investing in it at $t = 0$.

This creates complementarity between investment and advertising. Since investing in an advertised asset is safer, the arbitrageur will take a larger position in it than for an identical non-advertised asset. Formally, the arbitrageur's expected gain from investing in an advertised asset is a convex function of the mass of informed investors trading it, so that his gain from investing in it is greater if the attention of informed investors concentrates on K assets only. 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 on a single stock only.

3.3 Advertising heterogeneous assets

If assets are symmetric, as assumed so far, the arbitrageur is indifferent as to which to advertise. If instead assets have different characteristics, the arbitrageur will prefer to advertise some rather than others. Intuitively, he will choose to advertise the K most profitable assets. Recall that, from Proposition 1, if the arbitrageur already advertises K symmetric assets, then advertising an extra one is not profitable, even if this extra asset is as profitable as the original K assets. With heterogeneous assets, the extra asset is not going to be more profitable than the K assets already advertised, so that *a fortiori* advertising it will not be optimal.

Accordingly, in this case too the arbitrageur will advertise exactly K assets. To analyze which ones, we note that the equilibrium payoff from advertising asset i is given by expression (12), whereas the corresponding payoff from not advertising it is expression (13). Substituting for k_i from (11), the gain from advertising asset i is

$$u_i^A - u_i^N = \frac{(\hat{\theta}_i - P_0^i)^2}{2\rho} \frac{\tau_A^i(2\tau_S^i + \tau_A^i)}{\tau_S^i + \frac{{\tau_S^i}^2}{\tau_A^i} + \sigma_{u_i}^2 \rho^2}. \tag{15}$$

This equation immediately yields several comparative statics results:

Proposition 2 (**Characteristics of advertised assets**). *The gain from advertising asset i is increasing in its mispricing $(\hat{\theta}_i - P_0^i)^2$ and in the precision of the arbitrageur's private information τ_A^i, and is decreasing in the variance of noise trading $\sigma_{u_i}^2$.*

These results are intuitive: the arbitrageur will seek to advertise the asset with the highest expected return and lowest risk. Expected return is highest when mispricing is greatest, while risk is lowest for the arbitrageur when his private information is most precise (τ_A^i)

and his exposure to noise trading risk is lowest ($\sigma^2_{u_i}$). From the arbitrageur's standpoint, noise traders create additional risk at liquidation and so are undesirable. Hence, markets with many noise traders are unattractive for advertising arbitrageurs, and may exhibit large price deviations from fundamentals.

Proposition 2 highlights the marked difference between the advertising arbitrageur and a typical insider in the spirit of Kyle (1985). In that setting, noise trading increases the insider's profits by allowing the insider to camouflage his information. In contrast, in this model the arbitrageur wants to broadcast his information as widely as possible and regards noise traders as a nuisance. The key difference between the two settings is that in ours the arbitrageur anticipates that he will liquidate his position inelastically at $t = 1$, so that at that stage he can no longer extract any trading profits from his private information and is instead exposed to noise traders' risk: advertising then becomes the best way to funnel his private information indirectly into prices and reduce risk from noise trading. By contrast, in Kyle's model the insider always exploits his private information by trading against noise traders, thus benefiting from their presence.

3.4 Optimal portfolio

We can now fully characterize the optimal portfolio of the arbitrageur at $t = 0$. As argued in the previous section, he chooses to advertise K assets, whose characteristics are illustrated by Proposition 2. Then, expression (9) yields his optimal investment in the advertised asset i for $m = 1$, and in non-advertised assets j for $m = 0$:

$$x_i = \frac{\hat{\theta}_i - P_0^i}{\rho} \frac{(\tau_S^i + \tau_A^i)^2}{\tau_S^i + {\tau_S^i}^2/\tau_A^i + \sigma^2_{u_i}\rho^2}, \tag{16}$$

$$x_j = \frac{\hat{\theta}_j - P_0^j}{\rho} \frac{{\tau_S^j}^2}{\tau_S^j + {\tau_S^j}^2/\tau_A^j + \sigma^2_{u_j}\rho^2}. \tag{17}$$

The arbitrageur takes positions in all M assets he has private information about, even in the $M - K$ assets that he does not advertise. For instance, if $\hat{\theta}_j > P_0^j$ he will take a long position in asset j to exploit its underpricing, as shown by (17).

However, advertising has an additional effect on his portfolio choice and his profits, because it makes his position in the advertised asset less risky and therefore allows him to trade more aggressively on his private information. To distinguish this effect from that of private information *per se*, imagine that asset i is not advertised. In this case the arbitrageur's position in it would be given by (17): we denote this hypothetical position by $\tilde{x}_i$. Advertising would induce the arbitrageur to take an even more skewed position in asset i. Indeed, the ratio between his optimal position in asset i and the hypothetical position $\tilde{x}_i$ is greater than 1:

$$\frac{x_i}{\tilde{x}_i} = 1 + \Delta_i = \left(1 + \frac{\tau_A^i}{\tau_S^i}\right)^2 > 1, \tag{18}$$

where Δ_i measures the percentage increase in the position attributable to advertising.

As a result of his more aggressive position in the advertised asset i, on average the arbitrageur makes higher profits on that asset than on the non-advertised assets about which he has private information. His expected profits from trading these two assets are proportional to the respective optimal positions (16) and (17), so that the relative gain in profitability from advertising is also proportional to Δ_i: $\pi_i/\tilde{\pi}_i = 1 + \Delta_i$.

The fact that advertised assets yield extra profits is consistent with the evidence displayed by Ljungqvist and Qian (2016). Equation (18) highlights the main determinant of these extra profits:

Proposition 3 (Portfolio bias). *The percentage increase Δ_i in the arbitrageur's position in asset i and in the expected profits due to advertising this asset is increasing in the precision of his information relative to that of the public signal τ_A^i/τ_S^i.*

Intuitively, the arbitrageur chooses to overweight the advertised asset by Δ_i because he knows that by advertising it he will push the liquidation price P_1^i closer to the asset's fundamental value, which he knows, and thereby reduce the risk stemming from this unbalanced position. The magnitude of this effect increases with the relative precision of the arbitrageur's information τ_A^i/τ_S^i, because attentive investors, when trading at $t = 1$, will place more weight on his signal relative to the public information and accordingly push the liquidation price P_1^i closer to his estimate of the fundamental.

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 an individual 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 investors have the capacity to process the information on all these assets, i.e. $A \cdot M \leq K$, then trivially all arbitrageurs will advertise all the private information they have: in this case, the limit to information capacity is not binding. Here instead, we consider the more interesting case 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 a common scarce resource, i.e. investors' attention, just as in the tragedy of the commons: advertising too many assets would lead to too little attention being paid to each one 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. In principle, they may have "common information" about the same set or "exclusive information" about distinct sets of assets. Common information may arise either because arbitrageurs happen to receive the same signal independently, or through information sharing prior to advertising and trading. The common information case is the more interesting of the two, as it is the one where arbitrageurs are in principle able to solve the coordination problem: when they have superior information about the same assets, they may be able to agree informally on which ones to advertise. Hence, we focus the analysis mainly on this case (equilibrium under exclusive information is discussed in Section 4.3.

In most of the analysis, for simplicity we assume identical assets with independent returns, but briefly illustrate how the results would change with heterogeneous assets, differing, say, in extent of mispricing. This extension highlights another interesting margin on 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: arbitrageurs have information about the same set of $A \cdot M$ assets, possibly as a result of information sharing, and may agree to advertise $L \leq A \cdot M$ of them. Investors divide their attention equally among the advertised assets, so that the fraction of investors who are informed about any one is

$$m(L) = \min\left[1, K/L\right]. \tag{19}$$

Each arbitrageur i knows that L assets are being advertised, so that from expression (14) his expected payoff is

$$U_i(L) = kL\left[(\tau_S + \tau_A m(L))^2 - \tau_S^2\right] + AM\tau_S^2, \tag{20}$$

which coincides with the payoff of the single arbitrageur in expression (14), except that the last term is multiplied by A. Intuitively, the first term, capturing the incremental utility from advertising L assets, is the same as in the case of the single arbitrageur, while the last term, which captures the gains from trading on private information, is now larger because each arbitrageur draws on the common pool of information regarding all $A \cdot M$ assets. Because all arbitrageurs have the same information, and they coordinate, each will get the payoff in (20), and they will choose L so as to maximize it. Since only the first term depends on L, the number assets advertised, it is immediate that the optimal choice of L is the same as for the single arbitrageur. Hence, Proposition 1 applies: the total payoff of arbitrageurs is maximized when they just saturate the information processing capacity of investors by setting $L = K$, so that $m = 1$ and each arbitrageur gets the maximal expected payoff

$$U^* = kK \left[(\tau_S + \tau_A)^2 - \tau_S^2 \right] + AM\tau_S^2. \tag{21}$$

Hence, if arbitrageurs coordinate their decisions, the concentrated advertising principle of Proposition 1 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, the others 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 across them, multiple arbitrageurs will refrain from advertising an asset different from that advertised by others, to avoid distracting investors: each has the incentive to "piggyback" on others' advertising.

Whether the advertising is done by a single arbitrageur or by a coordinated group, arbitrageurs want investors to receive just the greatest amount of information they can process, so as to maximize the effect on prices. This enables all of them to take large positions in those assets, knowing that their private information is likely to affect the liquidation price. In either case, the complementarity between advertising and investment decisions is at the core of this concentration result.[10] Naturally, even if assets are heterogeneous, the choices of coordinated arbitrageurs are identical to those of a monopolistic arbitrageur: in both cases, they will want to exploit the best arbitrage opportunities available to them, e.g. advertise the most severely mispriced assets, as shown in Proposition 2.

This equilibrium behavior may appear to resemble the herding induced by information cascades, but in fact it is quite different: in this model, arbitrageurs all picking the same assets depends on common fundamental information and on strategic complementarity, not

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

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 Competition with Common Information

Now we turn to the more interesting case where arbitrageurs share information about all assets but fail to coordinate about which to advertise. Let us denote by $\mathbf{M}$ the set of $A \cdot M$ assets that arbitrageurs can advertise in this case. If advertising is costless, an arbitrageur may choose to advertise any subset $\mathbf{m}_i \subseteq \mathbf{M}$ of assets. Since investors can only process K signals, they waste no attention on multiple advertisements about the same asset, as each contains the same information. In other words, recognizing that any two messages are identical is assumed not to require attention. This implies that investors will pay attention to only one advertisement per asset, and so split their attention across the L unique advertisements in the set $\mathbf{L} = \bigcup_{j \in A} \mathbf{m}_j = \mathbf{L}_i \cup \mathbf{L}_{-i}$, 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.

Since investors split their attention randomly among $L_i + L_{-i}$ assets, the fraction of investors who are informed about any advertised asset is

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

In this case, the expected payoff of arbitrageur i becomes

$$U_i(L_i; L_{-i}) = k(L_i + L_{-i}) \left[(\tau_S + \tau_A m(L_i, L_{-i}))^2 - \tau_S^2 \right] + AM\tau_S^2, \tag{22}$$

where the first term captures arbitrageur i's incremental utility from his own advertising and from the advertising of others and the last term captures the gains from trading on all the shared information.

To show how many assets are advertised in the absence of coordination, consider first that in the benchmark case where arbitrageurs could coordinate on the optimal number to advertise, they would choose $L_i + L_{-i} = K$, as shown in the previous subsection. But now the equilibrium can differ from the coordinated outcome. In expression (22), two possible cases may arise. If other arbitrageurs are believed to advertise $L_{-i} < K$ assets, then arbitrageur i will find it optimal to advertise $K - L_{-i}$ additional assets, so as to precisely saturate investors' informational capacity. If instead other arbitrageurs already advertise $L_{-i} \geq K$ assets, then arbitrageur i will be indifferent about advertising any of the L_{-i} assets, but would never find it optimal to advertise an asset not included in L_{-i}, as this would lower his objective (22). Hence, there are multiple equilibria, and in some of them the number of assets advertised exceeds investors' processing capacity, i.e. $L_{-i} > K$, as stated in the next proposition.

Proposition 4. *With common information and competition, in equilibrium only* $L \in [K, A \cdot M]$ *assets can be advertised.*

By Proposition 1, it is immediate that when $L > K$ the equilibrium payoff of an investor is lower than the maximal payoff U^* obtained in the benchmark case, as defined by (21). In other words, too many assets are advertised in equilibrium, leading to an inefficient dispersion of investors' attention.

If the assets about which arbitrageurs are informed differ in their characteristics, competition entails another possible inefficiency, namely that the assets advertised are not those yielding the highest possible payoff to arbitrageurs. To illustrate this point, suppose that there are only two types of asset, which differ in the gain (15) that advertising them generates for arbitrageurs: advertising a "good" asset (G) yields a greater gain than advertising a "bad" one (B). We also assume that half of the assets the arbitrageur is informed about are good and the other half bad, and that $A \cdot M/2 > K$, so that the number of assets in each class is sufficient to saturate investors' capacity. For simplicity, suppose that there is no public signal about assets, so that $\tau_S = 0$ in (15) for both assets. Clearly, a monopolistic arbitrageur would always prefer to advertise a good rather than a bad asset, as the resulting payoff from expression (15) would be larger:

$$U_G \equiv \frac{(\hat{\theta}_G - P_0^G)^2 \tau_A^{G2}}{\sigma_u^{G2}} \geq \; U_B \equiv \frac{(\hat{\theta}_B - P_0^B)^2 \tau_A^{B2}}{\sigma_u^{B2}}. \tag{23}$$

For example, this condition is satisfied for the assets that are more severely mispriced, i.e. $\hat{\theta}^G - P_0^G > \hat{\theta}^B - P_0^B$, or for which arbitrageurs have more precise information, i.e. $\tau_A^G > \tau_B^G$, or there is less noise trading, i.e. $\sigma_u^{G2} < \sigma_u^{B2}$.

However, this may not be the case when multiple arbitrageurs decide which assets to advertise competitively; that is, asset heterogeneity may entail an additional source of inefficiency with multiple arbitrageurs. Indeed, if assets are not too widely different, i.e. if the following condition holds

$$U_G \leq U_B \left(2 + \frac{1}{AM}\right), \tag{24}$$

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 5. *With common information and competition, if condition (24) holds, then* *in equilibrium any combination of good and bad assets can be advertised, the total number of assets advertised is $L \in [K, AM]$.*

Intuitively, condition (24) guarantees that even if only $L \leq AM$ 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 better ones that would deliver as much as twice as large an 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 bad assets (for instance those with only mild mispricing) are advertised, even advertising a better (e.g. a more severely mispriced) one would reduce the fraction of investors informed about each and, due to convexity, disproportionately reduce their payoff.

In the Proof of Proposition 5 we show that, for any number $L \leq AM$ of assets advertised in equilibrium the no-deviation condition is weaker than (25), namely:

$$U_G \leq U_B \left(2 + \frac{1}{L}\right), \tag{25}$$

Intuitively, the no-deviation condition (25) becomes weaker as the number of assets advertised, L, increases, because for large L only a few investors pay attention to each advertised asset anyway, so the convexity of the arbitrageurs' payoff function plays little role (being locally close to linear).

This inefficiency is due to competition among arbitrageurs. In fact, it did not arise in Section 4.1 where arbitrageurs were assumed to be able to coordinate: in that setting, they would collectively choose to advertise the best assets only, namely, those with the largest mispricing $(\hat{\theta} - P_0)^2$, highest precision of private information τ_A and/or lowest noise trading σ_u^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 mispricings, if only they could coordinate their investment and advertising on those rather than on lesser ones, as we have shown in Section 4.1.

This also implies that arbitrageurs' trading behavior should be closely correlated, as in any equilibrium they have the incentive to trade advertised assets more intensively than other assets they are informed about. This is consistent with the evidence in Luo (2018) that around the date when a hedge fund pitches a stock at an investment conference, other hedge funds take positions similar to that of the pitching hedge fund, and they all liquidate these positions subsequently, like the arbitrageurs in our model. At the same time, Luo (2018) finds that mutual funds only started buying the stocks after they were pitched by hedge funds, thus behaving as rational attentive investors in our model.

These correlated trading strategies are also reminiscent of the "wolf pack" activism by hedge funds documented by Becht et al. (2017) and modeled by Brav et al. (2018). Just as in their model activists implicitly coordinate with many followers in engaging target management, in our equilibria informed arbitrageurs are predicted to trade the same advertised assets, resulting in highly correlated trading even though they act non-cooperatively. Of course, this results applies *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 a Nash equilibrium.

4.3 Exclusive Information

Now 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 effectively there is no information sharing or 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 arbitrageurs now never advertise the same assets, the fraction of informed investors about any advertised asset is

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

From (14), the expected payoff of arbitrageur i can be expressed as

$$U_i(L_i; L_{-i}) = kL_i\left[(\tau_S + \tau_A m(L_i, L_{-i}))^2 - \tau_S^2\right] + M\tau_S^2, \tag{27}$$

where the first term is the extra payoff from the L_i advertised assets and the second is the baseline payoff from trading on private information for all M assets. Unlike the common information case, here no benefit accrues to the arbitrageur from advertising by other arbitrageurs or information in common with them: accordingly, his payoff (27) differs from its analogues (22) and (20).

As a result, the equilibrium advertising with exclusive information also differs from those obtaining under common information, and indeed invariably leads to a lower payoff, as arbitrageurs always overload investors with advertising:

Proposition 6. *In an equilibrium with exclusive information, each arbitrageur advertises all M assets he is informed about, so that the total number of advertised assets $A \cdot M$ inefficiently exceeds the information processing capacity of investors 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 commitment to 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 have been assumed to be small relative to the market, so that they have no price impact. While this may be a natural assumption for a single arbitrageur, it becomes more questionable if many arbitrageurs trade the same asset. Consider for instance the case of an underpriced asset which a number of arbitrageurs bought at $t = 0$: when they collectively liquidate their position at $t = 1$, this sale will tend to lower the asset's price, and therefore their expected profits. This adverse price effect is the focus of this section.

To investigate this point, we consider the case in which arbitrageurs are able to coordinate their advertising activity as in Section 4.1: recall that in this case, absent any price impact, they would choose to advertise exactly K assets, so as not to overload investors' attention. In principle, with price impact the desire to mitigate the adverse effect of their trades may induce them to advertise more than K assets, so as to spread their trades across a larger number of assets. Hence, their search for greater liquidity might be expected to lead them to over-saturate investors' information-processing capacity. But we show here that this is not the case: even with price impact, arbitrageurs will coordinate on the same outcome that obtains without it; namely, they advertise only K assets.

As in Section 4.1, we consider A arbitrageurs with common information about $A \cdot M$ symmetric assets, i.e. with identical parameters τ_A, τ_S, $\hat{\theta}$, and P_0. We also assume that they are able to build up their position at $t = 0$ with no impact on P_0, while their trades do affect the liquidation price P_1: allowing P_0 too to react to their trades, while complicating the analysis, would not change the results qualitatively. Moreover, for simplicity we assume there is no public signal about assets, i.e. $\tau_S = 0$. Each arbitrageur decides which assets $i = 1, .., A \cdot M$ to advertise. Under these assumptions, we prove the following result:

Proposition 7. *Suppose that arbitrageurs take into account the price impact of their trades, are informed about $A \cdot M$ symmetric assets, and are able to coordinate their advertising strategies. Then, if investors can process at most K signals, in equilibrium only K assets are advertised.*

Hence, the result of concentrated advertising holds even if arbitrageurs are aware of their individual and collective impact on market prices: avoiding dilution of investors' attention dominates any gain from spreading the impact of their trades across multiple assets.

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" (Ljungqvist and Qian (2016), p. 2011).

(ii) Advertising accelerates price discovery, and on average it increases arbitrageurs' profits: this prediction is consistent with the finding of Ljungqvist and Qian (2016) that on average the price of the stocks targeted by the arbitrageurs in their sample drop by 7.5% on the date arbitrageurs release their first report, and by 21.4% to 26.2% in the three subsequent months, and with that of 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.

(iii) Advertising by arbitrageurs who are known to have precise private information has greater price impact. This prediction squares with the finding of Ljungqvist and Qian (2016) that prices react more strongly to reports by arbitrageurs whose previous recommendations have proved to be correct, and of Chen et al. (2014) that recommendations published by investors who correctly predicted past abnormal returns have a stronger price impact than those of other investors.

(iv) Different arbitrageurs will tend to advertise the same opportunities and to 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. Similarly, Luo (2018) reports evidence of correlated trading by hedge funds when one of them pitches a stock.

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

(i) Arbitrageurs should overweight advertised assets in their portfolios, benchmarked against the allocation that they choose when they do not advertise them, and such overweighting should be greater, the more precise the arbitrageurs' private information (as proxied, say, by their track record), and the more imprecise the public information about the asset.

(ii) Arbitrageurs are more likely to advertise assets that are more severely mispriced, those for which their private information is more precise, and those that are less exposed to noise trading shocks.

(iii) Advertising by arbitrageurs should peak when investors are not distracted by salient aggregate news, so that most of their attention capacity is available to process the advertised information.

(iv) The larger the number of competing arbitrageurs who simultaneously advertise different assets, the smaller the acceleration of price discovery for the corresponding asset and the profit for the respective arbitrageurs.

Appendix

Proof of Proposition 1. When the number of assets advertised rises from L to $L+1$, then the arbitrageur's expected payoff (14) changes by

$$\Delta U(L) = k\tau_A(L+1)m(L+1)(\tau_A m(L+1) + 2\tau_S) - k\tau_A L m(L)(\tau_A m(L) + 2\tau_S).$$

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

$$\Delta U(L) = k\tau_A(\tau_A + 2\tau_S).$$

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:

$$\Delta U(L) = -\frac{k\tau_A^2 K^2}{(L+1)L}.$$

Hence, if the arbitrageur has information about $M > K$ assets, then he will entirely use up investors' attention, but not over-exploit it. Obviously, if $M \leq K$, then he will advertise all M assets he is informed about.

QED.

Proof of Proposition 5. First, consider a candidate equilibrium where all arbitrageurs advertise the same $L \geq K$ bad assets, so that investors' attention capacity is already saturated by information about them (as required by Proposition 1). 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 payoff from advertising the L bad assets is

$$U_i = L\frac{(\hat{\theta}_B - P_0^B)^2}{2\rho\sigma_u^{B2}}\left(\tau_A^B\frac{K}{L}\right)^2.$$

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 payoff becomes

$$U_i' = \frac{(\hat{\theta}_G - P_0^G)^2}{2\rho\sigma_u^{G2}}\left(\tau_A^G\frac{K}{L+1}\right)^2 + L\frac{(\hat{\theta}_B - P_0^B)^2}{2\rho\sigma_u^{B2}}\left(\tau_A^B\frac{K}{L+1}\right)^2$$

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

$$U_i' - U_i = \frac{(\hat{\theta}_G - P_0^G)^2}{2\rho\sigma_u^{G2}}\left(\tau_A^G \frac{K}{L+1}\right)^2 - \frac{(\hat{\theta}_B - P_0^B)^2}{2\rho\sigma_u^{B2}}\left(\tau_A^B \frac{K}{L+1}\right)^2\left(2+\frac{1}{L}\right) \leq 0,$$

which, upon recalling the definition of good and bad assets in (23), becomes (25)

$$U_G \leq U_B\left(2+\frac{1}{L}\right). \tag{28}$$

We have shown that there is an equilibrium in which only bad assets are advertised if (25) holds, because no arbitrageur would prefer to deviate and advertise a good asset. Note that condition (25) implies (24) for any $L \leq AM$. Naturally, if arbitrageurs advertise any combination of good and bad assets in equilibrium, they get a higher payoff than from advertising bad assets only, which 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 6. It is easy to see that $U(L_i; L_{-i})$ given by (27) increases with L_i if $K > L_{-i} + L_i$, so that in equilibrium it cannot be $K > L_{-i} + L_i$, as in this case some arbitrageur i would deviate by increasing L_i. Next, notice that if $K < L_{-i} + L_i$, the derivative of $U(L_i, L_{-i})$ with respect to L_i is

$$\frac{\partial U(L_i; L_{-i})}{\partial L_i} = kK\frac{\tau_A}{L_{-i}+L_i}\left(2\tau_S\frac{L_i}{L_{-i}+L_i} + \tau_A\frac{K(L_{-i}-L_i)}{L_{-i}+L_i}\right).$$

In a symmetric equilibrium, $L_i = L^*$, $L_{-i} = (A-1)L^*$, which implies $L_{-i} > L_i$ and $\frac{\partial U(L_i;L_{-i})}{\partial L_i} > 0$. 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 7. To derive the equilibrium prices of the assets at $t = 1$, each arbitrageur $h = 1, .., A$ is assumed to acquire positions x_{ih}, $i = 1, .., L$, in L advertised assets, and to liquidate them, i.e. trade $-x_{ih}$, at $t = 1$. In deriving these initial positions, the advertising decisions of the arbitrageurs are taken as given. We denote by m_i the fraction of rational investors who have learned signal θ_i from the arbitrageurs' advertising. In total, the informed investors trade an amount $m_i y_I$, while the remaining fraction $1 - m_i$ of rational investors remain uninformed and trade $(1 - m_i)y_U$. Note that these investors are atomistic and their individual trades have no price impact, so that their individual trades are given by the same formulas as in Section 2.

The market clearing condition for asset i at $t = 1$ is

$$(1 - m_i)y_U + m_i y_I + u_i - \sum_{h=1}^{A} x_{ih} = 0,$$

and the resulting equilibrium price is

$$P_1^i = \hat{\theta} + \frac{\rho}{m_i \tau_A}(u_i - \sum_{h=1}^{A} x_{ih}). \tag{29}$$

As of time $t = 0$, the arbitrageur expects the price at $t = 1$ to be given by (29), which is distributed according to $P_1^i \sim N(\hat{P}_1^i, \sigma_P^2(m_i))$, where

$$\hat{P}_1^i = \mathbb{E}[P_1^i | \hat{\theta}, X_i] = \hat{\theta} - \frac{\rho}{m_i \tau_A} \sum_{h=1}^{A} x_{ih}, \tag{30}$$

and

$$\sigma_P^2(m_i) = \frac{\sigma_u^2 \rho^2}{m_i^2 \tau_A^2}. \tag{31}$$

Hence at $t = 0$ the optimal investment x_{ij} of the arbitrageur j solves

$$\max_{\{x_{ij}\}}[(\hat{\theta} - \frac{\rho}{m_i \tau_A}(x_{ij} + \sum_{h \neq j} x_{ih}) - P_0^i)x_{ij} - \frac{\rho}{2} x_{ij}^2 \sigma_P^2(m_i)], \tag{32}$$

where we substituted for $\hat{P}_1^i$ from (30). The investment x_{ih} chosen by other arbitrageurs ($h \neq j$) solves an analogous problem.

Notice that here the arbitrageur takes into account the price effect of his own trade and of the trades of the other arbitrageurs. As arbitrageurs are identical, and have the same information about the L assets, in equilibrium they choose the same investment $x_i = x_{ij}$ so as to satisfy the first-order condition of the maximization problem (32):

$$x_i = \frac{\hat{\theta} - P_0}{\rho} \frac{m_i^2 \tau_A^2}{(A+1)m_i \tau_A + \sigma_u^2 \rho^2}. \tag{33}$$

Comparing this expression with its analogue under no price impact in (9) shows that, as expected, arbitrageurs reduce their trades when they take their adverse price impact into account.

Substituting for x_i from (33) into the arbitrageur's objective function in (32) yields the expected gain from investing in asset i in equilibrium:

$$u(m_i) = \frac{(\hat{\theta} - P_0)^2}{\rho} \frac{m_i^2 \tau_A^2}{\left((A+1)m_i\tau_A + \sigma_u^2\rho^2\right)^2} \left(m_i\tau_A + \frac{\sigma_u^2\rho^2}{2}\right). \tag{34}$$

Building on expression (34), one can analyze arbitrageurs' optimal advertising decisions. If arbitrageurs advertise $L \leq K$, then the fraction of informed investors about each advertised asset is $m_i = 1$, and the expected payoff of each arbitrageur is

$$U_h(L) = L\frac{(\hat{\theta} - P_0)^2}{\rho} \frac{\tau_A^2}{\left((A+1)\tau_A + \sigma_u^2\rho^2\right)^2} \left(\tau_A + \frac{\sigma_u^2\rho^2}{2}\right).$$

Since for $L \leq K$ the payoff increases linearly in L, at least K assets will be advertised in equilibrium. When exactly K assets are advertised, the equilibrium payoff is

$$U_h(K) = K\frac{(\hat{\theta} - P_0)^2}{\rho} \frac{\tau_A^2}{\left((A+1)\tau_A + \sigma_u^2\rho^2\right)^2} \left(\tau_A + \frac{\sigma_u^2\rho^2}{2}\right).$$

In principle, in a symmetric equilibrium the arbitrageurs may choose to advertise more assets. When $L > K$ assets are advertised, then the fraction of investors informed about each asset is $m_i = K/L$, and the arbitrageur's payoff becomes

$$U_h(L) = L\frac{(\hat{\theta} - P_0)^2}{\rho} \frac{\tau_A^2 (K/L)^2}{\left((A+1)\tau_A K/L + \sigma_u^2\rho^2\right)^2} \left(\tau_A K/L + \frac{\sigma_u^2\rho^2}{2}\right)$$

Each arbitrageur prefers to advertise K assets if $U_h(K)/U_h(L) \geq 1$, which is equivalent to

$$\frac{L}{K}\left(\frac{(A+1)\tau_A K/L + \sigma_u^2\rho^2}{(A+1)\tau_A + \sigma_u^2\rho^2}\right)^2 \frac{\tau_A + \frac{\sigma_u^2\rho^2}{2}}{\tau_A K/L + \frac{\sigma_u^2\rho^2}{2}} \geq 1. \tag{35}$$

Rearranging, this condition becomes

$$\left((A+1)\tau_A + \frac{L}{K}\sigma_u^2\rho^2\right)^2 \left(\tau_A + \frac{\sigma_u^2\rho^2}{2}\right) - \left((A+1)\tau_A + \sigma_u^2\rho^2\right)^2 \left(\tau_A + \frac{L}{K}\frac{\sigma_u^2\rho^2}{2}\right) \geq 0.$$

This is equivalent to

$$\begin{aligned}\left((A+1)^2\tau_A^2 + \frac{L^2}{K^2}\sigma_u^4\rho^4 + 2(A+1)\tau_A\frac{L}{K}\sigma_u^2\rho^2\right)\left(\tau_A + \frac{\sigma_u^2\rho^2}{2}\right) - \\ \left((A+1)^2\tau_A^2 + \sigma_u^4\rho^4 + 2(A+1)\tau_A\sigma_u^2\rho^2\right)\left(\tau_A + \frac{L}{K}\frac{\sigma_u^2\rho^2}{2}\right) \geq 0,\end{aligned} \tag{36}$$

that is,

$$-(A{+}1)^2\tau_A^2\frac{L-K}{K}\frac{\sigma_u^2\rho^2}{2}{+}2(A{+}1)\sigma_u^2\rho^2\tau_A^2\frac{L-K}{K}{+}\sigma_u^4\rho^4\left(\tau_A\frac{L^2-K^2}{K^2}+\frac{L}{K}\frac{L-K}{K}\frac{\sigma_u^2\rho^2}{2}\right)\geq 0,$$

or equivalently

$$\frac{L-K}{K}\left(\frac{3}{2}(A+1)\sigma_u^2\rho^2\tau_A^2+\sigma_u^4\rho^4\left(\tau_A\frac{L+K}{K}+\frac{L}{K}\frac{\sigma_u^2\rho^2}{2}\right)\right)\geq 0,$$

which holds for any $L \geq K$. Hence, arbitrageurs never advertise more than K assets.

QED.

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.

Brav, A., A. Dasgupta, R. D. Z. Mathews, and P. Kumar (2018): "Wolf Pack Activism," *ECGI Finance Working Paper No. 501/2017*.

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.

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: The State of the Theory," *NBER Working paper*.

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.

Kyle, A. S. (1985): "Continuous Auctions and Insider Trading," *Econometrica*, 53, 1315–1335.

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

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*.

Pasquariello, P. and Y. Wang (2018): "Speculation with Information Disclosure," *SSRN Working paper*.

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

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*.