Advertising Arbitrage*

Sergei Kovbasyuk † Marco Pagano ‡

March 12, 2015

Abstract

Speculators often advertise arbitrage opportunities in order to persuade other investors and thus accelerate the correction of mispricing. We show that in order to minimize the risk and the cost of arbitrage an investor who identifies several mispriced assets optimally advertises only one of them, and overweights it in his portfolio; a risk-neutral arbitrageur invests only in this asset. The choice of the asset to be advertised depends not only on mispricing but also on its "advertisability" and accuracy of future news about it. When several arbitrageurs identify the same arbitrage opportunities, their decisions are strategic complements: they invest in the same asset and advertise it. Then, multiple equilibria may arise, some of which inefficient: arbitrageurs may correct small mispricings while failing to eliminate large ones. Finally, reputation matters: the ads of arbitrageurs who proved right in the past have a stronger price effect than the ads of arbitrageurs with no previous track record – and even stronger than those who proved wrong in the past.

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, Thierry Foucault, Mikhail Golosov, Antonio Guarino, Hugo Hopenhayn, Tullio Jappelli, Ralph Koijen, Guy Laroque, Marco Ottaviani, Nicola Pavoni, Nicola Persico, Andrea Pozzi, Wenlan Qian, Jean Tirole and especially to Alexander Ljungqvist for insightful remarks and suggestions. We also thank participants to seminars at Bocconi, CSEF, EIEF, Toulouse, UCL, Vienna, EFA and AEA annual meetings, and the 10th CSEF-IGIER Symposium on Economics and Institutions for their comments. We acknowledge financial support from EIEF.

†EIEF. 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 of them have taken to more than simply disclose their positions and expressing opinions, and back their thesis with data on allegedly mispriced assets. Examples range from such large hedge funds as David Einhorn's Greenlight Capital talking down and shortselling the shares of Allied Capital, Lehman Brothers and Green Mountain Coffee Roasters, to small investigative firms (like Muddy Waters Research, Glaucus Research Group, Citron Research and Gotham City Research) shorting companies, while providing evidence of fraudulent accounting and recommending "sell."[1] This advertising activity is associated with abnormal returns: Ljungqvist and Qian (2014) examine the reports that 17 professional investors published upon shorting 113 US listed companies between 2006 and 2011, and find that they managed to earn substantial excess returns on their short positions, especially when their reports contained hard information. Similar evidence arises 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 thus eventually on stock prices.

These examples tell a common story: some investors who detect mispriced securities (hereafter, "arbitrageurs") 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 if the advertising is successful it will nudge prices closer to fundamentals, and enable the arbitrageurs to close their positions profitably. This mechanism is crucially important for arbitrageurs who are too small to influence prices by their own trading; to muster the requisite fire-power they need to bring other investors to their side. This is the case of the investors studied by Ljungqvist and Qian (2014) and by Chen et al. (2014), who are are so small and constrained that they cannot hope to correct the mispricing just by trading the targeted stocks.

However profitable on average, this business practice is both costly and risky: uncovering and advertising hard information is costly and, once the information is divulged, other investors may disregard it, because they are either inattentive or unconvinced. In this case, stock prices will fail to react to the arbitrageur's advertising effort or even move adversely to his position, inflicting losses on him, as vividly illustrated by this recent episode:

[1]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.

> "At a crowded hall in Manhattan, Bill Ackman, an activist hedge-fund manager, at last laid out his case alleging that Herbalife is a pyramid scheme. Mr Ackman has bet $1 billion shorting Herbalife's shares and spent $50m investigating its marketing practices. During his presentation he compared the company to Enron and Nazis, but the 'death blow' he said he would deliver failed to pack a punch; Herbalife's share price rose by 25% by the end of the day." (The Economist, 26 July 2014).

In this paper we show that these costs and risks have several non-trivial implications for the portfolio choices of arbitrageurs that engage in advertising, the intensity of their advertising activity, and its impact on securities' prices.

First, even when an arbitrageur identifies several mispriced assets, he will concentrate his advertising on a single one: drawing the attention of other investors to a single asset, he is most likely to eliminate its mispricing, while dispersing the advertising effort across several assets would likely fail to end mispricing in any. That is, concentrated advertising is a safer bet than diversified advertising: it increases the chances that the arbitrageur will close his position profitably.

Second, concentrating advertising on a single asset produces portfolio under-diversification. Advertising a mispriced asset raises the short-term payoff and lowers the short-term risk, so even a risk-averse arbitrageur will want to overweight the asset that he advertises, and a risk-neutral one will hold only that asset.

Third, in order to save on advertising costs and maximize the return on their position, arbitrageurs will prefer the most "advertisable" and most mispriced assets among those that they may target, and advertise such assets most intensively. Hence, simple and familiar assets are more likely to be targeted by arbitrageurs and intensively advertised than complex and unfamiliar ones. Arbitrageurs are also more likely to invest in assets for which they expect precise public information to emerge in the future, as this allows them to save on advertising costs: the price of such assets will converge to its fundamental value even without much advertising.

Fourth, again to save on advertising costs, arbitrageurs will tend to advertise the same asset as others: by advertising an asset, each arbitrageur makes it more profitable for others to invest in it as well; and once they are exposed to the risk from this asset, the other arbitrageurs will want to advertise it. However, mutual "piggybacking" by arbitrageurs tends to generate multiple equilibria, some of which are inefficient: arbitrageurs may be collectively trapped in an inefficient portfolio choice, where they all advertise an asset that is not the most seriously mispriced. Indeed, if there are enough arbitrageurs, they may end up collectively picking any of the mispriced assets, even the least underpriced. This may explain why the market sometimes appears to pick up the minor mispricing of some assets, and neglect the much more pronounced mispricing of others, especially complex ones like RMBSs and CDOs before the subprime financial crisis.

Finally, a solid reputation may allow an arbitrageur to save on advertising costs, or equivalently make a given advertising effort more effective. An arbitrageur with a good reputation may be able to publicize his recommendations even if he does not justify them with hard data: the price reaction to his advertising is proportional to his reputation. As arbitrageurs build a good track record, the price reaction to their advertising intensifies, but if their recommendations turned out to be wrong this effect disappears. In fact, the data analyzed by Ljungqvist and Qian show that arbitrageurs move prices more sharply when they can show a history of credible advertising. This may also explain why the market often heeds the recommendations of well-known investors even when they are not backed by solid evidence: for instance, on 13 August 2013, on Icahn's buy recommendation on Twitter, the price of Apple rose by 5%.

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), 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: it succeeds precisely when it catches the attention of investors, i.e. when it induces them to devote their scarce processing ability to the opportunity identified.[2] Advertising also adds a dimension that is missing 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 model arbitrageurs choose under-diversified portfolios, like investors in Van Nieuwerburgh and Veldkamp (2009, 2010), but for a different reason. Our arbitrageurs have unlimited information-processing capacity (and are perfectly 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. Second, our arbitrageurs' herd behavior is superficially reminiscent of

[2]The same result would obtain if information about mispricing were costly to acquire, rather than hard to process: in this case advertising would work by conveying information to investors free of charge rather than directing their attention to it. So our model can be reinterpreted as based on costly information acquisition.

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. In 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.

The result that arbitrageurs can develop reputation and move prices with soft information would appear to parallel Benabou and Laroque (1992), who show that market gurus can affect prices even if they are believed to be honest only on average.[3] In both models, arbitrageurs' or gurus' track record affects their credibility. But in our model advertising is never deliberately deceptive: some arbitrageurs can successfully forecast future returns, others can't; the advertisements of the latter are likely to be misleading, but not purposely so. In contrast, in Benabou and Laroque (1992) gurus have perfect information, but sometimes are dishonest, lying to investors in order to make profits.

Finally, our analysis of the interactions among arbitrageurs can be related to Abreu and Brunnermeier (2002), who argue that arbitrage may be delayed by synchronization risk: in their model, arbitrageurs learn about an arbitrage opportunity sequentially, and thus prefer to wait when they are unsure that enough of them have learnt about 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, mispricing is known to all arbitrageurs, so there is no synchronization risk, but advertising may lead them to coordinate on the "wrong" asset. Hence, while advertising does mitigate the limits to arbitrage, it may not remove them altogether, insofar as the collective behavior of the arbitrageurs may not touch on the most acute mispricing.

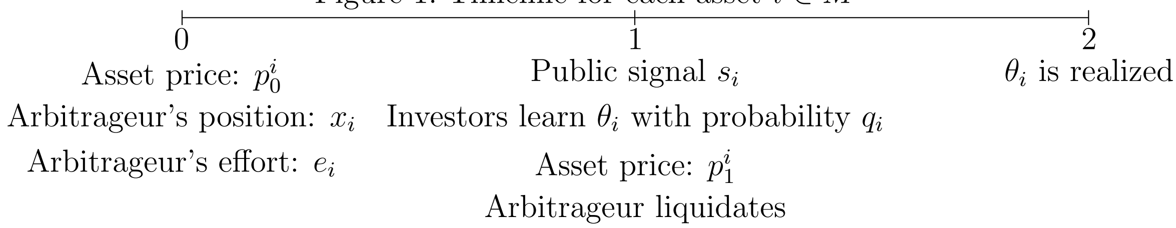
The paper is organized as follows. Section 1 introduces the model. Section 2 characterizes the arbitrageur's advertising. In section 3 we study how advertising affects the portfolio choice of risk-averse arbitrageurs. Section 4 examines the case of risk-neutral arbitrageurs and how asset characteristics affect both advertising and portfolio choices. Section 5 allows for strategic interactions among arbitrageurs. Section 6 investigates the way in which arbitrageurs' reputation affects the effectiveness of their advertising, in a setting where their reputation evolves over time as a function of their past performance. The last section summarizes and discusses our predictions.

[3] In their model, the guru's information cannot be justified with hard evidence. Instead, the guru is believed to be honest with a given probability and to be opportunistic with the complementary probability. If the guru is opportunistic and gets positive private information about the asset, he sends a negative message that drives the price down, buys cheap and gets a high return. Benabou and Laroque conclude that if they have some reputational capital gurus can manipulate markets.

1 Environment

The baseline model has a single arbitrageur in a market of many risk-neutral investors. There are three periods: $t = 0, 1, 2$, and there is a continuum of assets ($i \in N$), traded at dates $t = 0, 1$ and delivering return $\theta_i \in \{0, 1\}$ at $t = 2$. At $t = 0$ investors' prior belief about the return is given by $\Pr(\theta_i = 1) = \pi_i$, where for technical reasons we assume $\pi_i \in [\underline{\pi}, \overline{\pi}]$, $i \in N$, $0 < \underline{\pi} < \overline{\pi} < 1$. Investors have no discounting. The timeline is as follows:

Figure 1: Timeline for each asset $i \in M$



At $t = 0$, the arbitrageur privately learns θ_i for a finite subset of assets $i \in M$ and decides in which assets to take positions.[4] In principle, the arbitrageur can take a position x_i in any asset $i \in N$. Assets that do not belong to the set M are of no interest for the arbitrageur because he has no private information about them; hence, without loss of generality we consider assets in M.

At $t = 0$ asset i can be traded at price p_0^i: the arbitrageur takes position x_i and decides on advertising effort e_i, $i \in M$. At $t = 1$ with probability $q_i(e_i)$ the arbitrageur's advertising is successful and investors learn θ_i; with complementary probability, the advertising is not effective and investors rely on noisy public signal s_i about θ_i. The signal is correct ($s_i = \theta_i$) with probability $\gamma_i \in [0, 1)$ and is an uninformative random variable ϵ_i with probability $1 - \gamma_i$. Its distribution is the same as that of θ_i: $\Pr(\epsilon_i = 1) = \pi_i$, $i \in N$, but it is independent of θ_i. This random variable can be seen as arising from one of two sources: mistakes in public announcements or noise trading. Therefore, γ_i affects the signal-to-noise ratio of the price at $t = 1$.

At $t = 1$ each asset i can be traded at p_1^i, so that the arbitrageur's monetary payoff is $c = \sum_i x_i p_1^i$. Finally, at $t = 2$ all assets' final returns θ_i, $i \in M$ are realized. The arbitrageur cannot wait until the final returns are realized, so he liquidates his portfolio at $t = 1$. This captures the urgency of either investing in other profitable assets or consuming. 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.

[4] Investors are assumed not to know the set M: they believe that any asset $i \in N$ is in M with the same probability. Otherwise, information about assets in M may also be relevant for assets outside M.

The arbitrageur's utility $V(c)$ is an increasing function of his monetary payoff at $t = 1$, $c = \sum_i x_i p_1^i$. We assume that the utility function is not convex, that is, we allow the arbitrageur to be either risk-averse or risk-neutral: formally, $V'(c) > 0$, $V''(c) \leq 0$.

At $t = 0$ the arbitrageur has monetary resources $w > 0$ that he allocates among investments x_i. Denoting by $y_i = |x_i p_i^0|$ the absolute market value of the arbitrageur's position in asset i at $t = 0$, his budget constraint is

$$\sum_{i \in M} y_i \leq w. \tag{1}$$

Notice that (2) also imposes a constraint on the arbitrageur's short positions, because in practice both long and short positions require some collateral.

At $t = 0$, the arbitrageur also chooses his advertising effort $e_i \geq 0$, $i \in M$. This effort is necessary to prove and expose the mispricing: it includes the collection of hard evidence and its dissemination.The arbitrageur can devote a maximum amount of effort $E > 0$ to produce and communicate convincing evidence about θ_i, $i \in M$:[5]

$$\sum_{i \in M} e_i \leq E. \tag{2}$$

Investors have limited attention, in the sense that they can learn θ_i only if an arbitrageur advertises asset i. And even so, advertising is not necessarily successful:

Assumption 1. ***(Investors' limited attention)*** *Investors learn the true realization of θ_i at $t = 1$ only if advertising is effective, which happens with probability $q_i(e_i)$ for any asset $i \in M$. The advertising function is weakly increasing: $q_i'(e_i) \geq 0$ for $e_i \in [0, E]$, $q_i(0) = 0$ and $q_i(E) = \overline{q}_i \in (0, 1)$ for any $i \in M$.*

Hence, the arbitrageur can never induce investors to learn his information for sure, even if he devotes all of his effort to a single asset: $q_i(E) = \overline{q}_i < 1$. The parameter $\overline{q}_i$ captures the extent to which information about asset i is "advertisable": $\overline{q}_i$ may be high when investors are very receptive to information about asset i, for instance because they already hold it, or because it belongs to a relatively well-known class. Investors' attention may also be affected by the asset's previous performance – how often, say, the asset has been in the news previously.

If advertising fails, investors learn the true θ_i only at $t = 2$. Investors do not worry about potential market manipulation by arbitrageurs: here we posit that arbitrageurs advertise hard information. Section 6 explores advertising based on soft information.

[5]We introduce a separate effort budget for simplicity: in principle we can allow E to also include money.

We assume that advertising features no economies of scope, in the sense that distributing the total effort E among several assets does not make advertising more effective than concentrating all of it on one of the assets. Formally:

Assumption 2. ***(No economies of scope in advertising)*** *For any* $(e_1, ...e_M)$ *such that* $\sum_{i \in M} e_i = E$*, the advertising functions* $q_i(e_i)$ *satisfy:*

$$\sum_{i \in M} \frac{q_i(e_i)}{\overline{q}_i} \leq 1. \tag{3}$$

To gain intuition about this assumption, suppose that the arbitrageur has information only about two assets i and j with the same advertisability ($\overline{q}_i = \overline{q}_j = \overline{q}$). Then, Assumption 2 implies that $q_i(e_i) + q_j(E - e_i) \leq \overline{q}$: if the arbitrageur were to split his total effort E between the two assets, he would not be more likely to convey his information about either asset to investors than if he were to devote his entire effort only to one of the two. In contrast, economies of scope in advertising would imply $q_i(e_i) + q_j(E - e_i) > \overline{q}$ for $e_i \in (0, E)$, so that an arbitrageur who splits his effort between the two assets is more likely to convey his information about at least one of them to investors.

A simple example in which Assumption 2 holds is that of linear advertising functions:

$$q_i(e_i) = \frac{\overline{q}_i}{E} e_i, \ i \in M. \tag{4}$$

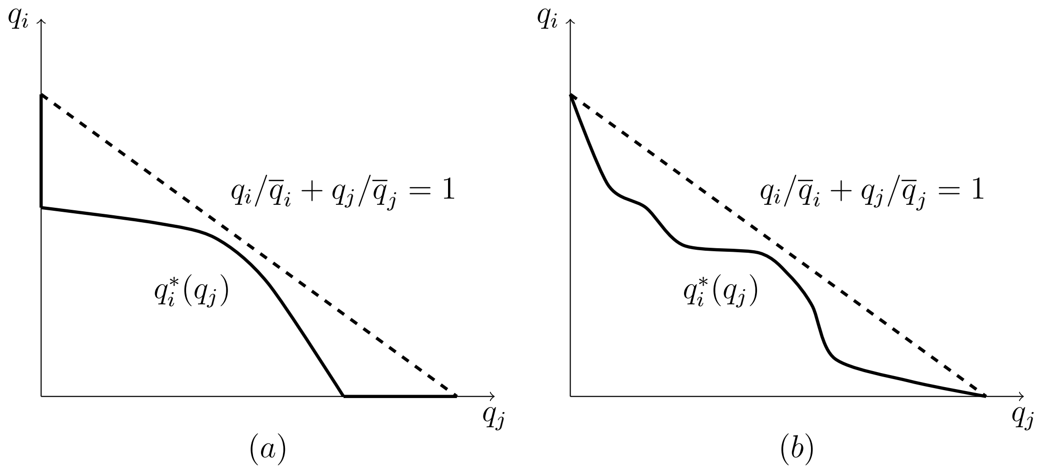
In this case the weak inequality in Assumption 2 holds with equality, namely:

$$\sum_{i \in M} \frac{q_i(e_i)}{\overline{q}_i} = \sum_{i \in M} \frac{e_i}{E} = 1, \text{ s.t. } \sum_i e_i = E.$$

If we consider only two assets i and j, linear advertising functions imply that feasible values of q_i and q_j satisfy $q_i/\overline{q}_i + q_j/\overline{q}_j = 1$, i.e. lie on the dashed line in Figures 2 (a) and (b).

Beside this linear case, Assumption 2 encompasses all advertising functions such that the set of feasible values (q_i, q_j) lies below the dashed lines in Figure 2. The solid lines in Figure 2 depict the frontier $q_i^*(q_j)$ of the feasible set in two natural examples. The first, shown in Figure 2 (a), is the case where advertising requires both a fixed effort cost and a variable one: a minimum amount of effort is necessary to gather any evidence about the asset. The other, in Figure 2 (b) is the case where the marginal productivity of advertising is low for small advertising effort. The example in Figure 2 (b) captures the argument made by the literature on the value of information, starting with Radner and Stiglitz (1984), according

Figure 2: Feasible values q_i and q_j under different assumptions.



to which small amounts of information have zero value to a decision maker (see also Keppo at el. (2008) for a recent and general analysis): in our context, outside investors rationally ignore a small amount of evidence produced by the arbitrageur about any given asset. The common feature of these two cases us that little effort is ineffective in convincing investors, i.e. $q_i(e_i) = 0$ for small e_i.

We assume that the arbitrageur's financial resources w are not only limited but small, in the sense that his trades are negligible against the total market volume of any asset: he acts as a price taker. The arbitrageur can affect asset prices only by advertising his private information:

Assumption 3. ***(Price taking)*** *Arbitrageur's trades do not affect prices.*

For brevity, and with no loss of generality, we consider only undervalued assets:

Assumption 4. ***(Undervalued assets)*** *The true value of any asset $i \in M$ is $\theta_i = 1$.*

Clearly, the arbitrageur may only want to take long positions in these assets $x_i \geq 0$, $i \in M$. All results also hold if we assume $\theta_i = 0$ in M and study short positions.

2 Concentrated advertising

We now solve for the arbitrageur's advertising effort and portfolio choice. At $t = 0$ the risk-neutral investors have prior beliefs π_i about asset $i \in M$, such that the price is $p_0^i = \pi_i$. At $t = 1$ investors learn θ_i with probability q_i, in which case the price becomes $p_1^i = \theta_i$. With complementary probability $1 - q_i$, investors do not learn θ_i and rely only on the public signal s_i; in this case the price is $p_1^i = E[\theta_i|s_i] = (1 - \gamma_i)\pi_i + \gamma_i s_i$. The signal s_i

is correct with probability γ_i, and the prior about θ_i is π_i, so that by Bayesian updating investors' expectation is $E[\theta_i|s_i] = (1-\gamma_i)E[\theta_i|\epsilon_i = s_i] + \gamma_i E[\theta_i|\theta_i = s_i] = (1-\gamma_i)\pi_i + \gamma_i s_i$.

The return from investing in the asset at $t = 0$ is $\tilde{r}_i = \frac{p_1^i}{p_0^i}$, with three possible values: $r_i^H = \frac{1}{\pi_i}$ if advertising succeeds; $r_i^M = 1 - \gamma_i + \frac{\gamma_i}{\pi_i}$ if advertising fails and $s_i = 1$; $r_i^L = 1 - \gamma_i$ if advertising fails and $s_i = 0$.

At $t = 0$ the arbitrageur knows $\theta_i = 1$, for $i \in M$. From the arbitrageur's standpoint $\Pr(s_i = 1|\theta_i = 1) = \gamma_i \Pr[\theta_i = 1|\theta_i = 1] + (1-\gamma_i)\Pr[\epsilon_i = 1|\theta_i = 1] = \gamma_i + (1-\gamma_i)\pi_i$. For brevity we denote $t_i = \Pr(s_i = 1|\theta_i = 1)$ and $1 - t_i = \Pr(s_i = 0|\theta_i = 1)$, $i \in M$. The distribution of asset i's return to the arbitrageur is

$$\tilde{r}_i = \begin{cases} r_i^H & \text{with probability } q_i \\ r_i^M & \text{with probability } (1-q_i)t_i \\ r_i^L & \text{with probability } (1-q_i)(1-t_i) \end{cases}, i \in N. \tag{5}$$

The arbitrageur chooses his portfolio holdings $\mathbf{y} = (y_1, ..., y_M)$ and his advertising efforts $\mathbf{e} = (e_1, ..., e_M)$ at $t = 0$. At $t = 1$ his final wealth is $c = \sum_{i=1}^M \tilde{r}_i y_i$. For instance, if the arbitrageur were to advertise all assets and investors were to learn all θ_i, $i \in M$ at $t = 1$, the arbitrageur's monetary payoff would be $c = \sum_{i=1}^M r_i^H y_i$, which happens with probability $\prod_{i \in M} q_i$.

At $t = 0$ the arbitrageur maximizes his expected utility taking (2) and (5) into account. The return on each asset $i \in M$ has three possible realizations; thus for two assets we have nine possible realizations of the monetary payoff, and for M assets we have 3^M possible realizations. In general, the expression for expected utility is very cumbersome. For conciseness, we pick any two assets i and $j \neq i$ from M, and consider four states of advertising effectiveness: (i) successful for both i and j, (ii) successful only for i, (iii) successful only for j and (iv) not successful either for i or j. If the advertising of asset i is not successful, its return can be described by a binary random variable $\rho_i \in \{r_i^M, r_i^L\}$, with $\Pr(\rho = r_i^M) = t_i$. Analogously for j. The returns of all assets $\tilde{r}_i$, $i \in M$, are independent. For brevity, denote by $\tilde{r}_{-ij} = \sum\limits_{k \neq i,j} \tilde{r}_k y_k$ the return on other assets in M except i and j. Then we can write the arbitrageur's expected utility at $t = 0$ as follows:

$$\begin{aligned} &E[V|\mathbf{y}, \mathbf{e}] = q_i q_j E[V(y_i r_i^H + y_j r_j^H + \tilde{r}_{-ij})] + q_i(1-q_j)E[V(y_i r_i^H + y_j \rho_j + \tilde{r}_{-ij})] + \\ &(1-q_i)q_j E[V(y_i \rho_i + y_j r_j^H + \tilde{r}_{-ij})] + (1-q_i)(1-q_j)E[V(y_i \rho_i + y_j \rho_j + \tilde{r}_{-ij})]. \end{aligned} \tag{6}$$

The portfolio and advertising decisions maximize the arbitrageur's expected payoff (6) taking into account the budget constraint (1) and the effort feasibility constraint (2):

$$\max_{\{\mathbf{y}\geq 0, \mathbf{e}\geq 0\}} E[V|\mathbf{y}, \mathbf{e}], \ s.t. \ \sum_i y_i \leq w, \ \sum_i e_i \leq E. \tag{7}$$

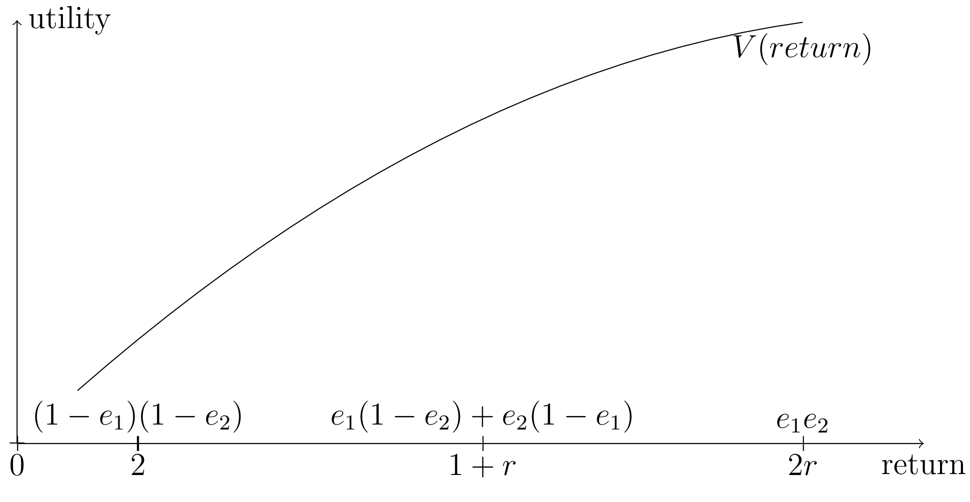
The first step to solve the arbitrageur's problem is to characterize his advertising decisions, taking his investment decisions as given:

Proposition 1. *The arbitrageur advertises only one asset:* $e_i > 0$ *for some* $i \in M$ *and* $e_j = 0$ *for any* $j \neq i$.

All proofs are in the appendix. The proof of Proposition 1 is straightforward if the arbitrageur is risk-neutral. Intuitively, a risk-neutral arbitrageur invests in the asset with the highest expected return, and advertises an asset only if he invests in it. Therefore, a risk-neutral arbitrageur does not advertise two assets. But if the arbitrageur is risk-averse, the result is not obvious. One may imagine that in this case the arbitrageur would choose to buy and advertise several assets in order to diversify risk. But this is not true. We illustrate the intuition with a simple symmetric example with two assets and no informative public signal (for the detailed proof see the appendix).

Example with two assets. M contains two identical assets $i = 1, 2$ such that $\gamma_1 = \gamma_2 = 0$, $r_1^L = r_2^L = 1$, $r_1^H = r_1^M = r_2^H = r_2^M = r > 1$, $\overline{q}_1 = \overline{q}_2 = 1$. For the sake of illustration, suppose that the arbitrageur has no cost of effort but a single unit of advertising capacity, which he can either allocate equally to both assets ($e_1 = e_2 = 1/2$) or concentrate entirely on one of them ($e_i = 1$, $e_{-i} = 0$, $i = 1, 2$). Also, suppose that $w = 2$ and the arbitrageur invests $y_1 = y_2 = 1$ in each asset. We can show that advertising both assets delivers a lower expected payoff than advertising only one.

Figure 3: Symmetric example with two assets.



Suppose the arbitrageur advertises both assets: $e_1 = e_2 = 1/2$. With probability $(1 - e_1)(1 - e_2) = 1/4$, his advertising is ineffective for both assets, and his monetary payoff is $y_1 r_1^L + y_2 r_2^L = 2$; with probability 1/4, advertising is effective for both assets and the monetary payoff is $y_1 r_1^H + y_2 r_2^H = 2r$; with probability 1/2, advertising is effective for only one asset, and the monetary payoff is $1 + r$. The arbitrageur's expected utility is thus $E[V|e_1 = \frac{1}{2}, e_2 = \frac{1}{2}] = \frac{1}{4}V(2) + \frac{1}{4}V(2r) + \frac{1}{2}V(1 + r)$.

Suppose instead that the arbitrageur advertises only one asset, setting for instance $e_1 = 1$, $e_2 = 0$. With probability $e_1 = 1$, his advertising on asset 1 is successful, while that on asset 2 is never effective. Hence, with certainty he gets return $1 + r$ and his expected utility is $E[V|e_1 = 1, e_2 = 0] = V(1 + r)$.

The difference in payoffs is $E[V|e_1 = 1, e_2 = 0] - E[V|e_1 = \frac{1}{2}, e_2 = \frac{1}{2}] = \frac{1}{2}V(1 + r) - \frac{1}{4}V(2) - \frac{1}{4}V(2r)$. Since the arbitrageur is risk-averse, we have $V(1+r) > \frac{1}{2}V(2) + \frac{1}{2}V(2r)$, that is, the arbitrageur prefers to advertise only one asset. This apparently counter-intuitive result is actually very natural. Advertising both assets produces a riskier lottery than advertising only one, because the former is a mean-preserving spread of the latter. Hence, the risk-averse arbitrageur prefers the latter, and advertises only one asset.

Let us describe the general intuition behind this result. A risk-averse arbitrageur tries to insure against a bad outcome of no information at $t = 1$ by advertising and thus increasing the probability of information arriving at $t = 1$. For a given portfolio choice, he prefers to allocate all his advertising effort to a single asset, precisely because he is risk-averse and concentrating advertising on one asset is a safer bet than spreading it across several assets. When he advertises a single asset, it is most likely that at time $t = 1$ this asset will deliver a high return, while the other, unadvertised assets are most likely not to deliver high returns. As a result, the payoff involves little risk. But if he were to spread advertising effort across assets, many would pay off with some probability and the final payoff would be very uncertain. This parallels the choice of the "job market paper" in the academic market: typically, candidates come to the market with a single strong paper. Betting your career on a single paper may seem to be a highly risky strategy, but our analysis suggests that it is actually the safest, allowing the candidate to devote all his or her energies on advertising a single project and gaining the market's attention.

3 Overweighting of the advertised asset

Proposition 1 greatly simplifies the analysis. As only one asset $i \in M$ is advertised, $q_i = \overline{q}_i$, $q_j = 0$ for $j \neq i$ and the expression for the arbitrageur's utility (6) can be written as

$$E[V|\mathbf{y}, e_i = E] = \overline{q}_i E[V(y_i r_i^H + \sum_{j \neq i} \rho_j y_j)] + (1 - \overline{q}_i) E[V(\sum_{j \in M} \rho_j y_j)]. \tag{8}$$

The arbitrageur's optimization problem (7) can be solved as follows. For each $i \in M$ and $q_i = \overline{q}_i$, find a portfolio $\mathbf{y}_i$ that maximizes (8) subject to $\sum_i y_i \leq w$, and denote by $E[V]_i$ the corresponding maximal value of the expected payoff. Choose the asset with the highest $E[V]_i$ among those in M. At the optimum the arbitrageur advertises asset $i^* \in \arg\max_{j \in M} E[V]_j$, and there may be multiple assets that deliver the same maximal payoff. The level of advertising is $e_{i^*} = E$ and the portfolio choice is $\mathbf{y}_{i^*}$.

As the above argument illustrates, once the arbitrageur has chosen the asset i^* and his advertising effort $e^*_{i^*} = E$, his portfolio choice becomes a standard diversification problem. The only difference is that the likelihood of a high return on asset i^* is enhanced by advertising. In general, one expects the arbitrageur to take a large position in asset i^* and small positions in the other assets, in order to reduce the overall riskiness of his portfolio. To show this point most clearly, we concentrate on a symmetric case where, in the absence of advertising, the arbitrageur would choose a balanced (equal-weighted) portfolio. With advertising, instead, he will overweight the advertised asset.

Assumption 5. *(Symmetric assets) Assets in M differ only in terms of advertisability: $\gamma_i = \gamma$ and $\pi_i = \pi$ for all $i \in M$ and $\overline{q}_i \neq \overline{q}_j$ for any $i \neq j$.*

As a benchmark case we solve for optimal portfolio allocation when advertising is not possible, i.e. $e = 0$. In this case all assets in M are equivalent: $t_i = t$, $r_i^M = r^M$, $r_i^L = r^L$ for all $i \in M$.

Lemma 1. *When advertising is not possible, the arbitrageur is risk-averse, and Assumption 5 holds, the arbitrageur takes equal positions in all assets in M.*

The lemma is intuitive. Given that assets have identical and independently distributed returns, a risk-averse arbitrageur fully diversifies, taking equal positions in all assets in M. But when advertising is possible, this is not the case:

Proposition 2. *When advertising is possible, the arbitrageur is risk-averse, and Assumption 5 holds, the arbitrageur advertises the most advertisable asset and invests more in it than in any other asset: for $i = \arg\max_{j \in M} \overline{q}_j$ we have $y_i > y_j$ for any $j \neq i$. Investments in other assets are the same: $y_j = y$ for $j \neq i$.*

To see this, recall that, by Proposition 1, only one asset is advertised; and this is the most advertisable asset, which has the highest probability of successful advertising for the maximum advertising effort. Proposition 2 states that for this reason the arbitrageur overweights this asset in his portfolio.

Propositions 1 and 2 establish that the arbitrageur's advertising and investment will be concentrated under a general utility function V. To go one step further and explicitly characterize the asset that the arbitrageur chooses to advertise in terms of its potential return, quality of public signal and advertisability, we take the case of a risk-neutral arbitrageur. This specification will turn out to be useful also for subsequent extensions of the model.

4 Risk-neutral arbitrageur

From now on, we assume that the arbitrageur is risk-neutral with respect to his monetary payoff: $V(c) = c$. However, we drop Assumption 5 about asset symmetry and consider M assets with different expected returns ($1/\pi_i \neq 1/\pi_j$), different informativeness of the public signal ($\gamma_i \neq \gamma_j$), and different advertisability ($\overline{q}_i \neq \overline{q}_j$ for any $i \neq j$).

According to Proposition 1, a risk-neutral arbitrageur advertises only one asset (for convenience, asset i, that is $e_i > 0$). He also invests all his wealth w in this asset. To understand why, first observe that the arbitrageur is risk-neutral; so he only cares about the expected return, not about risk. Second, suppose he invests in a second asset, $j \neq i$, that he does not advertise: $e_j = 0$. This would be consistent with optimality if the expected returns of both assets were equal; otherwise, the arbitrageur would strictly prefer one of the two. But if the unadvertised asset j yields the same return as the advertised asset i, it would necessarily produce an even higher return if advertised. Hence, the arbitrageur will benefit by advertising asset j instead of asset i: by choosing $e'_j = e_i > 0$, $e'_i = 0$ and $y'_j = w$ he increases the expected return of asset j. This contradicts the initial assumption that it is optimal to invest in both assets. Therefore, the arbitrageur not only advertises one asset but also invests all his wealth in that asset.

Let the asset in which the arbitrageur invests all his wealth be asset k ($y_k = w$). As advertising succeeds with probability $\overline{q}_k$, the expected payoff from investing in asset k and advertising it is:

$$E[V|k] = w\left[\frac{1}{\pi_k} - (1-\gamma_k^2)(1-\overline{q}_k)\left(\frac{1}{\pi_k}-1\right)\right]. \quad (9)$$

Proposition 3. *The arbitrageur invests $y_i^* = w$ in asset $i = \arg\max_{k \in M} E[V]k$ and advertises it: other things being equal, he prefers an asset that is more advertisable (high $\overline{q}_k$), more significantly mispriced (high $1/\pi_k$), and with more precise public information (high γ_k).*

The proof is straightforward, one only needs to take the derivative of $E[V|k]$ with respect to corresponding characteristics. All three characteristics (potential return $\frac{1}{\pi_k}$, advertisability $\overline{q}_k$, and quality of public information γ_k) are desirable from the arbitrageur's point of view: As a consequence the choice of the investment asset involves a trade-off. For instance, the arbitrageur may be indifferent between an asset with high advertisability $\overline{q}_i$ and low potential return $1/\pi_i$, and one with low advertisability $\overline{q}_k$ and high potential return $1/\pi_k$.

To sum up, the more advertisable and the more mispriced an asset, the more likely it is to be targeted by arbitrageurs and intensely advertised by them. In contrast, the precision of public information increases the chances that an asset is targeted by arbitrageurs but reduces their advertising effort. This is because advertising effort is a costly substitute for public information: ex ante, arbitrageurs prefer assets with precise public information because it allows them to save on advertising costs; but, given the choice of an investment asset, they will advertise it more intensively if it features poor rather than precise public information.

5 Multiple arbitrageurs

When several arbitrageurs acquire private information about different assets independently, each of them behaves as described in the previous sections. But the analysis changes considerably if several arbitrageurs have private information about the same set of assets.

Consider $L \geq 2$ identical arbitrageurs that at $t = 0$ have the same information about a set of mispriced assets $M \in N$. After learning the actual θ_i for assets in M at $t = 0$, each arbitrageur $l \in M$ chooses his investments $\mathbf{y}_l$ and advertising efforts $\mathbf{e}_l$, taking the behavior of other arbitrageurs as given.

It is essential to specify how the joint advertising activity of arbitrageurs affects the beliefs of other investors. For tractability we assume that the advertising function is linear in the aggregate advertising effort:

$$q_i = \min\left[\frac{\overline{q}_i}{E}\sum_{l=1,..,L} e_i^l, 1\right], \tag{10}$$

which is the aggregate equivalent of (4), where all arbitrageurs are taken to have the same feasible effort E. As in the previous section, assets are assumed to have different potential returns $1/\pi_i$, $i \in M$.

The possible realizations of the return on investment in asset i are characterized by equation (5), as before. When arbitrageurs choose their investments and advertising efforts, they have common information about the set of assets M: hence the game among arbitrageurs is one of complete information. We look for a Nash equilibrium in pure strategies $(\mathbf{y}_l^*, \mathbf{e}_l^*)$, $l = 1, ..., L$. First, we show that in equilibrium all arbitrageurs invest in the same asset. Second, we determine which assets are advertised in equilibrium. Finally, we show that the equilibrium can be "inefficient": the arbitrageurs would be better off if they all invested in a different asset and advertised it.

Lemma 2. *In equilibrium all L arbitrageurs invest in the same asset.*

The proof is intuitive: in equilibrium, arbitrageurs cannot invest in different assets. If some arbitrageurs invest in asset j and others in asset $k \neq j$, then the expected return of both assets must be the same. This implies that for one of the assets, say asset k, we must have $q_k < 1$. (Indeed, if $q_k = q_j = 1$ then the corresponding expected returns $1/\pi_k$and $1/\pi_j$ would not be equal, a contradiction: recall that by assumption $1/\pi_i$ differ for $i \in M$). Consider now an arbitrageur who advertises asset j. If he deviated by investing all his wealth in asset k and advertising it, the expected return of asset k would increase, and the arbitrageur would benefit, which is a contradiction. It follows that in equilibrium all arbitrageurs must invest in the same asset.

Next, we want to identify the asset that arbitrageurs choose in equilibrium. Denote by i the asset that an arbitrageur would choose in isolation (absent other arbitrageurs):

$$i = \arg\max_{k \in M} E[V|k], \tag{11}$$

where $E[V|k]$ is defined by (9). Clearly, in the presence of multiple arbitrageurs it is an equilibrium if they all pick asset i: their collective advertising of asset i makes this choice even more attractive than if this asset were advertised only by one of them.

But other equilibria may also exist, in which arbitrageurs invest in an asset j different from i. By Lemma 2 in such an equilibrium they all invest in asset j and therefore, with no loss of generality, also spend all their effort in advertising this asset, so that $q_j = \min[\overline{q}_j L, 1]$. The following proposition states the condition for the existence of such an equilibrium:

Proposition 4. *: Any asset $j \in M$ can be advertised in equilibrium if*

$$\begin{aligned} &\frac{1}{\pi_j} - (1-\gamma_j^2)(1-\min[\overline{q}_j L, 1]) \left(\frac{1}{\pi_j} - 1\right) \geq \\ &\frac{1}{\pi_i} - (1-\gamma_i^2)(1-\overline{q}_i) \left(\frac{1}{\pi_i} - 1\right), \end{aligned} \tag{12}$$

where asset i is defined by (11).

Hence, if condition (12) holds for $j \neq i$, there are multiple equilibria. The proof is immediate: the expected payoff of each arbitrageur can be expressed as

$$V_j(L) = w\left[\frac{1}{\pi_j} - (1-\gamma_j^2)(1-\min[\overline{q}_j L, 1])\left(\frac{1}{\pi_j} - 1\right)\right]. \tag{13}$$

An equilibrium in which all arbitrageurs invest in asset j and advertise it exists if and only if $V_j(L) \geq E[V|i]$, that is, every arbitrageur prefers to invest in asset j, which is already advertised by the other $L-1$ arbitrageurs, rather than deviating by investing in asset i and advertising it. The condition $V_j(L) \geq E[V|i]$ yields inequality (12).

Intuitively, each arbitrageur has the incentive to "piggyback" on the advertising of others. Since an asset advertised by others is more likely to pay off at $t = 1$, any arbitrageur will be more willing to invest in it. But if the arbitrageur invests in the asset, he also has the incentive to advertise it because he is exposed to its risk. This equilibrium outcome may 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 asset is based on common fundamental information and on strategic complementarity, not on an attempt to gather useful information from the others' decisions: indeed, their correlated behavior speeds up price discovery, rather than delaying it as in cascades models.

When arbitrageurs are numerous, the strategic complementarity between them is strong: as a result, in equilibrium they may all concentrate their investment and advertising efforts on some asset $j \in M$, even if it is only moderately mispriced (low $1/\pi_j$) and difficult to advertise (low $\overline{q}_j$) compared to asset i that they would choose in isolation. This can happen for several assets, which establishes the multiplicity of equilibria.

Interestingly, some of these equilibria are inefficient, in the sense that arbitrageurs could benefit if they could coordinate on advertising a different asset. To see this most clearly, suppose that arbitrageurs are so numerous that if they all advertise any asset $j \in M$ their collective advertising will succeed for sure: $L\overline{q}_j \geq 1$ for any $j \in M$. In this case the condition (12) simplifies: any asset $j \in M$ can be advertised in equilibrium if

$$\frac{1}{\pi_j} \geq \frac{1}{\pi_i} - (1-\gamma_i^2)(1-\overline{q}_i)\left(\frac{1}{\pi_i} - 1\right). \tag{14}$$

The expected payoff in each of these equilibria is $V_j(L) = w/\pi_j$, i.e. initial wealth multiplied by asset j's expected return. Clearly, in an efficient equilibrium the arbitrageurs advertise the most mispriced asset, namely that with the highest expected return $1/\pi_j$. Yet, the arbitrageurs may neglect this asset and advertise a less mispriced asset it equilibrium.

Hence, the strategic complementarity between arbitrageurs may explain why financial markets sometimes focus on minor mispricing of some assets while neglecting much more significant mispricing of other assets, such as RMBSs, CDOs or Greek public debt before the recent financial crises. So, the strategic complementarity just highlighted provides a new explanation for the persistence of substantial mispricing, which differs from those 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 large mispricings, if only they could coordinate their investment and advertising on such mispricings rather than on lesser ones.

6 Reputation

Until now we have posited an arbitrageur with perfect information about asset returns θ_i, $i \in M$. On this assumption, when the advertising of asset i succeeds, investors learn θ_i and the price adjusts accordingly: $p_i = \theta_i$. Now we consider what happens if the arbitrageurs' advertisement may be inaccurate, so that the price reaction to their advertisement depends on their reputation. Since the arbitrageur's choice among different assets has already been studied earlier in the paper, in this section we simplify the analysis and assume that the arbitrageur has private information about a single asset i. We also assume that there is no public signal at $t = 1$: allowing for the public signal would not alter the qualitative results, but would complicate the algebra considerably.

The signal $\hat{\theta}_i$ that the arbitrageur observes about asset $i \in N$ may be imperfect depending on the arbitrageur's type $\tau \in \{L, H\}$. If he is high-skill ($\tau = H$), which happens with probability μ, the signal is perfect; if he is low-skill ($\tau = L$), which happens with probability $1 - \mu$, the signal is pure noise. That is, if the arbitrageur is high-skill the signal equals the true value θ_i; if he is low skill it is an independent and identically distributed variable $\psi_i \in \{0, 1\}$, with $\Pr(\psi_i = 1) = \pi$. As previously, we focus on the case where the realization of the signal is positive ($\hat{\theta}_i = 1$), so that the arbitrageur takes a long position: in the opposite case, the analysis is symmetric.

Only the arbitrageur knows his type: investors' prior belief about his skill is $\mu = \Pr(H)$. Hence, τ stands for the arbitrageur's ability to identify arbitrage opportunities and the corresponding evidence, while μ stands for his reputation on this score. Note that even the low-skill arbitrageur, who has no private information, may choose to advertise his signal $\hat{\theta}_i$ when his reputation allows him to affect prices ($\mu > 0$).

To explore how advertising is affected by reputation and how reputation evolves depending on performance, we extend the time line of the model, as shown in Figure 4. At $t = 0$ the arbitrageur learns his type $\tau \in \{L, H\}$ and observes $\hat{\theta}_i = 1$ for asset i. He can buy it at price p_0^i, and he chooses his advertising effort $e_i^\tau \in [0, E]$. At $t = 1$ investors observe the signal $\hat{s}_i = \hat{\theta}_i$ sent by the arbitrageur if his advertising was successful, which happens with probability $q_i(e_i^\tau)$; with complementary probability they do not observe it, so that $\hat{s}_i = \emptyset$. Given $\hat{s}_i$, investors form beliefs about the arbitrageur's type $\mu(\hat{s}_i)$ and expectations about returns $E[\theta_i|\hat{s}_i]$, $i \in N$. Assets trade at prices p_1^i, $i \in N$, and the arbitrageur liquidates his position. At $t = 2$ assets produce their returns. Then, the whole sequence of actions just described is repeated. At $t = 3$ a new set N of assets with uncertain returns appears. The arbitrageur has a new endowment w, observes a signal about an asset $j \in N$, and decides on new investment and advertising. At $t = 4$ the advertising either succeeds or not, asset prices adjust, the arbitrageur liquidates and consumes. At $t = 5$ the new returns are realized.

Figure 4: Timeline in case of reputation building.

0	1	2	3	4	5
Type: τ	Signal: $\hat{s}_i$	θ_i occurs	Endowment: w	Signal: $\hat{s}_j$	θ_j occurs
Endowment: w	Price: p_1^i		Information: $\hat{\theta}_j$	Price: p_1^j	
Information: $\hat{\theta}_i$	Liquidation		Price: p_3^j	Liquidation	
Price: p_0^i			Choice of x_j, e_j^τ		
Choice of x_i, e_i^τ					

For simplicity, we assume that if advertising fails at $t = 1$ ($\hat{s}_i = \emptyset$), i.e. if investors neglect the arbitrageur's signal $\hat{\theta}_i$, the signal $\hat{\theta}_i$ that they missed cannot subsequently be retrieved.[6] Most of the results are qualitatively similar if we allow investors to go back and retrieve the arbitrageur's past signals if his later advertising succeeds and attracts their attention.

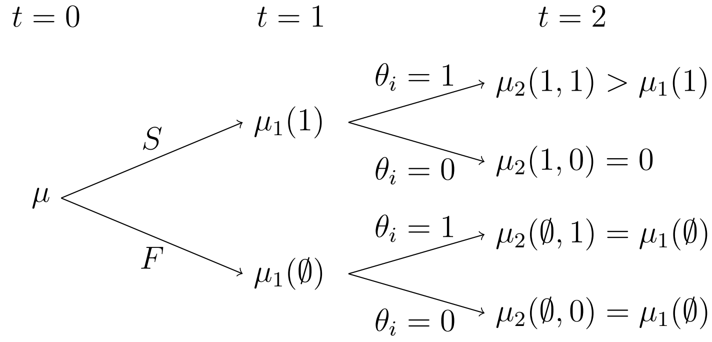
We solve for the Perfect Bayesian Equilibrium of this game. The equilibrium is characterized by the arbitrageur's decisions at $t = \{0, 3\}$ and investors' beliefs at $t = 0, .., 5$ about the arbitrageur's type $\mu_t(h_t)$ depending on the history h_t. The beliefs must be consistent with the arbitrageur's equilibrium strategy on the equilibrium path.

we first describe how investors' beliefs evolve for $t = 0, 1, 2$. At $t = 0$ investors believe that the arbitrageur type is high with prior probability μ. At $t = 1$ the investors' belief $\mu_1(\hat{s}_i)$ depends on whether advertising succeeds ($\hat{s}_i = \hat{\theta}_i = 1$) or not ($\hat{s}_i = \emptyset$). Finally, at

[6] Alternatively, one could assume that investors can process only one signal at a time: if so, it is not important if past signals are recorded or not because they will never choose to process an old signal if they can access a new one.

$t = 2$ the belief $\mu_2(\hat{s}_i, \theta_i)$ also depends on the realized return $\theta_i \in \{0, 1\}$. Figure 5 below summarizes the evolution of investors' beliefs.

Figure 5: Evolution of investors' beliefs



Let us start by deriving the posterior $\mu_2(\hat{s}_i, \theta_i)$ at $t = 2$ while taking posterior beliefs $\mu_1(1)$ and $\mu_1(\emptyset)$ at $t = 1$ as given (later we will characterize these beliefs as well). To describe the posterior belief $\mu_2(\hat{s}_i, \theta_i)$, note first that if advertising fails ($\hat{s}_i = \emptyset$), investors do not observe $\hat{\theta}_i$ and cannot compare it with the actual realization of θ_i, so that in this case they do not update their beliefs at $t = 2$: $\mu_2(\emptyset, 1) = \mu_2(\emptyset, 0) = \mu_1(\emptyset)$. Recall that at $t = 0$ the arbitrageur advertises signal $\hat{\theta}_i = 1$. If advertising is successful, investors observe $\hat{s}_i = \hat{\theta}_i = 1$ at $t = 1$. If at $t = 2$ the actual return is low ($\theta_i = 0$), then according to Bayes' rule the belief drops to zero: $\mu_2(1, 0) = 0$. Indeed, since the high-skill arbitrageur has perfect information $\hat{\theta}_i = \theta_i$, only the low-skill arbitrageur can advertise $\hat{\theta}_i = 1$ when actually $\theta_i = 0$. If instead advertising is successful ($\hat{s}_i = \hat{\theta}_i = 1$) and at $t = 2$ the actual return is high ($\theta_i = 1$), then the investors' belief becomes

$$\mu_2(1, 1) = \frac{\mu_1(1)}{\mu_1(1) + (1 - \mu_1(1))\pi} > \mu_1(1). \tag{15}$$

Since the high-skill arbitrageur always rightly identifies an undervalued asset ($\hat{\theta}_i = \theta_i$), the numerator of (15) is the joint probability of the arbitrageur being high-skill and $\hat{\theta}_i = \theta_i$. The denominator is the total probability of $\hat{\theta}_i = \theta_i$, because even the signal of the low-skill arbitrageur $\hat{\theta}_i$ with probability π coincides with θ_i. The fact that $\mu_2(1, 1) > \mu_1(1)$ indicates that when investors observe that the arbitrageur's advertisement was correct, they revise their belief upward.

At $t = 3$, the investors' belief about the arbitrageur's type $\mu_3(h_3)$ only depends on past history, since investors get no new information at $t = 3$. Hence their belief is the same as at $t = 2$: $\mu_3(h_3) = \mu_2(h_2) = \mu_2$. This allows us to describe how asset prices are affected by advertising occurring at $t = 3$. At $t = 3$ asset j trades at price $p_3^j = \pi_j$. At $t = 4$ new information arrives $\hat{s}_j \in \{\emptyset, 1\}$ and the asset's price is its expected value conditional on the signal and on the arbitrageur's reputation, i.e. the investors' posterior belief $\mu_4(\hat{s}_j)$ about his accuracy:

$$p_4^i = E_{\mu_4}[\theta_i|\hat{s}_i], \ i \in N. \tag{16}$$

Depending on investors' beliefs, there are two possible equilibrium outcomes: one without and one with advertising. If investors have pessimistic beliefs about credibility, then in equilibrium there is no advertising. Investors set $\mu_4(\hat{s}_i) = 0$ if they receive the arbitrageur's signal $\hat{s}_j = \hat{\theta}_j$, so that advertising has no effect on prices, and the arbitrageur does not find it optimal to advertise, irrespective of his type.

We focus on an equilibrium in which the arbitrageur advertises at $t = 3$, his signal raises the price at $t = 4$ (so that $p_4^j \geq p_3^j$), and the arbitrageur, being risk-neutral, invests his entire wealth in asset j ($x = w/\pi_j$) at $t = 3$. High-skill and low-skill arbitrageurs choose their advertising efforts e_j^H and e_j^L at $t = 3$ optimally, given investors' beliefs.

Consider first how investors update their beliefs at period $t = 4$: two cases are possible, depending on whether advertising fails ($\hat{s}_j = \emptyset$) or succeeds ($\hat{s}_j = 1$). If advertising fails, the investor's posterior belief about the arbitrageur's skill is

$$\mu_4(\emptyset) = \frac{\mu_2(1 - q(e_j^H))}{\mu_2(1 - q(e_j^H)) + (1 - \mu_2)(1 - q(e_j^L))}, \tag{17}$$

where the numerator is the joint probability of the arbitrageur being high-skill and not succeeding in advertising, while the denominator is the total probability of his advertising not being successful. Since in this case investors do not get any new information, they will not update their prior belief on asset values. Hence the price of asset i stays unchanged at its initial level:

$$p_4^j(\emptyset) = E[\theta_j] = \pi_j. \tag{18}$$

If advertising succeeds (so that $\hat{s}_j = \hat{\theta}_j = 1$), the investor's posterior belief about the arbitrageur's skill is

$$\mu_4(1) = \frac{\mu_2 q(e_j^H)}{\mu_2 q(e_j^H) + (1 - \mu_2) q(e_j^L)}, \tag{19}$$

where the numerator is the joint probability of the arbitrageur being high-skill and succeeding in advertising, and the denominator is the total probability of his advertising being successful. In this case, the price of asset j at $t = 4$ is

$$p_4^j(1) = \mu_4(1) + (1 - \mu_4(1))\pi_j, \tag{20}$$

i.e. the probability of the arbitrageur being high-skill multiplied by the true value of the asset, which in this case equals 1, plus the probability of his being low-skill multiplied by the prior valuation π_j.

The price reaction to advertising is the difference Δ between the two expressions just obtained for prices when advertising succeeds and when it fails:

$$\Delta \equiv p_4^j(1) - p_4^j(\emptyset) = \mu_4(1)(1-\pi_j). \tag{21}$$

At $t=3$ the arbitrageur takes this price reaction into account when he decides on advertising effort e_j. In expectation, asset j's price at $t=4$ is $E[p_4^j|e_j] = q(e_j)p_4^j(1) + (1 - q(e_j))p_4^j(\emptyset) = \pi_j + q(e_j)\Delta$, which can be expressed as $E[p_4^j|e] = \pi_j + q(e_j)\mu_4(1)(1-\pi_j)$ by using equations (18) and (20). The arbitrageur's expected payoff as of $t=3$ is $E[p_4^j|e_j]x^j + (w - x^j p_3^j)$, where x^j denotes the units of asset j that he buys at $t=4$. Since $p_3^j = \pi_j$ at $t=3$, this expected payoff equals $q(e_j)x^j\Delta + w$, which is weakly increasing in x^j and e_j because $\mu_4(1) \geq 0$. Hence in an equilibrium with advertising the following result holds:

Lemma 3. *The arbitrageur invests his entire wealth w in asset j ($x = w/\pi_j$) and chooses maximal advertising effort $e_j = E$ independently of his type.*

Using this result in (19) and (17), it immediately follows that $\mu_4(1) = \mu_4(\emptyset) = \mu_2$: as the two types of arbitrageurs advertise equally, their success or failure in advertising at $t=4$ is uninformative about their skill, and thus leads to no updating. Moreover, given the investors' belief μ_2, the expected payoff of an arbitrageur of any type as of $t=3$ can be expressed as

$$V'(\mu_2) = \overline{q}_j\mu_2(\frac{1}{\pi_j} - 1)w + w. \tag{22}$$

Clearly, when the arbitrageur decides on advertising for the first time at $t=0$, he anticipates how his effort will affect his future reputation μ_2 and his continuation payoff V' at $t=3$. At $t=1$ investors update their beliefs about the arbitrageur as described in Figure 5. The expressions for $\mu_1(\hat{s}_i)$ depending on success or failure of advertising at $t=1$ are analogous to (19) and (17), and those for the prices at $t=1$ are analogous to (20) and (18). Hence, the price reaction to advertising at $t=1$ is $\Delta_1 = \mu_1(1)(1-\pi_i)$ as in (21. This allows us to express the expected payoff of an arbitrageur of type τ at date $t=0$ as follows:

$$q(e_i^\tau)\frac{\Delta_1}{\pi_i}w + w + E[V'|e_i^\tau, \tau], \ \tau \in \{L, H\}, \tag{23}$$

where the last term takes into account the effect of investors' beliefs on the arbitrageur's future payoff V' after $t = 3$. Using Figure 5 we can express this continuation payoff as

$$E[V'|e_i^H, H] = q(e_i^H)V'(\mu_2(1,1)) + (1 - q(e_i^H))V'(\mu_1(\emptyset)) \tag{24}$$

for high-skill arbitrageurs, and as

$$E[V'|e_i^L, L] = q(e_i^L)[\pi_i V'(\mu_2(1,1)) + (1-\pi_i)V'(\mu_2(1,0))] + (1 - q(e_i^L))V'(\mu_1(\emptyset)) \tag{25}$$

for low skill ones.

Also at $t = 0$, we focus on equilibria in which at least one type of arbitrageurs advertises (again, an equilibrium with no advertising also exists). It is easy to see that the expected payoff of the high-skill arbitrageur is weakly increasing in e_i^H, so that in an equilibrium with advertising he chooses $e_i^H = E$. Consider a candidate equilibrium in which the low-skill arbitrageur also chooses $e_i^L = E$. In such an equilibrium we have $\mu_1(1) = \mu_1(\emptyset) = \mu$, $\Delta_1 = \mu(1 - \pi_i)$, and the expected payoff of a low-skill arbitrageur at $t = 0$ becomes:

$$q(e_i^L)\mu(\frac{1}{\pi_i} - 1)w + q(e_i^L)\pi_i \overline{q}_j \frac{\mu}{\mu + (1-\mu)\pi_i}(\frac{1}{\pi_j} - 1)w + (1 - q(e_i^L))\overline{q}_j \mu(\frac{1}{\pi_j} - 1)w + 2w.$$

This expression is increasing in e_i^L if and only if

$$\frac{1}{\pi_i} - 1 \geq \overline{q}_j \frac{\mu(1-\pi_i)}{\mu + (1-\mu)\pi_i}(\frac{1}{\pi_j} - 1),$$

which in turn always holds if $1/\pi_i \geq 1/\pi_j$, i.e. if the mispricing in period 0 is at least as large as in period 3. Since an equilibrium with $e_i^L = E$ exists if the expected payoff of a low-skill arbitrageur is weakly increasing in e_i^L, this establishes the following result:

Lemma 4. *An equilibrium in which both types of arbitrageur put maximal advertising effort at* $t = \{0, 3\}$ *exists if* $1/\pi_i \geq 1/\pi_j$.

Intuitively, the low-skill arbitrageur will advertise asset i at $t = 0$ if he expects to gain at least as much from that asset as from asset j at $t = 3$, because he knows that by time $t = 3$ he may have burnt his reputation. From now on we focus on this equilibrium, in which both types of arbitrageurs exert the maximal advertising effort $e_k^\tau = E$ for $k = i, j$ and $\tau = L, H$. In this equilibrium, the following result holds:

Proposition 5. *At* $t = 4$ *the price reaction to advertising is greatest when successful advertising at* $t = 1$ *correctly predicts the asset's return at* $t = 2$*, intermediate if advertising fails, and least if successful advertising at* $t = 1$ *is belied by the asset's return at* $t = 2$.

Hence, the price impact of advertising increases over time if in the past it proved to be accurate, while it drops if it proved wrong. The price reaction to advertising is determined by the reputation of the arbitrageur: whenever the reputation of a successful advertiser is good, the price reaction is also large.

7 Conclusions

Our model generates several testable hypotheses about the investment and advertising activity of arbitrageurs. Some still await empirical testing:

(i) Arbitrageurs concentrate advertising on one asset at a time: we should not find arbitrageurs advertising a new opportunity before cashing out on the previous one.

(ii) Arbitrageurs should overweight advertised assets in their portfolios, benchmarked against the portfolio allocation that they choose when they do not advertise them.

(iii) Arbitrageurs are more likely to advertise an asset – and to do so more intensively – if it is more severely mispriced and/or more advertisable than others. They will also advertise an asset more heavily when public information on it is less accurate (for instance, stocks that are not covered by analysts).

Others, however, have already been shown to be consistent with the evidence available:

(i) Advertising accelerates price discovery, and on average it increases arbitrageurs' profits: this prediction is consistent with the finding of Ljungqvist and Qian (2014), that on average the price of the stocks targeted by the arbitrageurs in their sample drop by 7.4% on the date arbitrageurs release their first report, and by 26.4% in the three subsequent months.

(ii) Advertising of hard information and advertising by reputable arbitrageurs has greater price impact. Both of these predictions are confirmed by Ljungqvist and Qian (2014), who show that reports based on actual data have a strong price impact, while those that contain only opinions have no significant effect, and that prices react more strongly to reports by arbitrageurs whose previous recommendations have proved to be correct. Similarly, Chen et al. (2014) document that recommendations published by investors who correctly predicted past abnormal returns have a stronger price impact than reports of other investors.

(iii) Different arbitrageurs will tend to advertise the same opportunities and to exploit them simultaneously. 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 by 324 to 376 basis points per month over the next two years.

Appendix

Proof of Proposition 1. The proof proceeds in two steps: i) then we relax the technology constraint (3) and find optimal adversing efforts in the relaxed problem, iii) we show that adversing efforts optimal in the relaxed problem are feasible in the initial problem and, hence, are optimal in the initial problem.

i). The choice of adversing efforts $\mathbf{e}$ uniquely defines the vector of probabilities $\mathbf{q} = (q_1, ..., q_M)$. All possible choices of $\mathbf{e} \geq \mathbf{0}$ such that $\sum_i e_i \leq E$ define the set of feasible probabilities $\mathbf{Q}$. Hence, with no loss of generality we can consider that the arbitrageur directly chooses $\mathbf{q} \in \mathbf{Q}$ which also satisfies the technology constraint (3): $\sum_i q_i/\overline{q}_i \leq 1$.

Consider a relaxed problem where the arbitrageur can choose any $\mathbf{q} \geq 0$, which satisfies $\sum_i q_i/\overline{q}_i \leq 1$, that is $\mathbf{q}$ may be outside $\mathbf{Q}$.

Note, that the arbitrageur's expected payoff is weakly increasing in q_i for any $i \in M$: $\frac{\partial E[V|\mathbf{y},\mathbf{e}]}{\partial q_i} = E[V(y_i r_i^H + \tilde{r}_{-i})] - E[V(y_i \rho_i + \tilde{r}_{-i})] \geq 0$ because $\rho_i \leq r_i^M < r_i^H$ and $y_i \geq 0$ (for brevity $\tilde{r}_{-i}$ denotes the return on all assets except i).

Hence, the highest expected payoff in the relaxed problem is achieved on the boundary of the feasible set, and we can consider $\sum_i q_i/\overline{q}_i = 1$ with no loss of generality. This constraint is equivalent to constraint $\sum e_i = E$ when advertising functions are linear $q_i = e_i \overline{q}_i/E$, $i \in M$. From now on we use this analogy. Since $\sum e_i = E$ we must have $e_i^* > 0$ for some $i \in M$ in optimum.

Notice that if the arbitrageur advertises asset i, he must have invested in it. Indeed if $y_j^* = 0$ then optimally $e_j^* = 0$, $j = 1, ..., M$; therefore $e_i^* > 0$ implies $y_i^* > 0$. Suppose there exists $j \neq i$ such that $e_j^* > 0$. This implies $y_j^* > 0$. Let $\hat{e} = e_i^* + e_j^*$, consider e_i and e_j such that $e_j = \hat{e} - e_i$ (note that these deviations satisfy $\sum e_i = E$).

According to Assumption 1 $q_i < 1$ and $q_j < 1$. A necessary condition for the maximum of the arbitrageur's expected payoff is that e_i and e_j maximize $E[V|\mathbf{y},\mathbf{e}]$ subject to $e_j = \hat{e} - e_i$. Use $q_k = e_k \overline{q}_k/E$ for $k = i, j$ and substitute $e_j = \hat{e} - e_i$ in (6). Suppose $e_j^* > 0$. The first order condition for an interior solution requires $\frac{\partial E(V|\mathbf{y},\mathbf{e})}{\partial e_i}|_{e_j = \hat{e} - e_i} = 0$.

We now show that this is not a maximum, and that an interior solution with $e_i > 0$ and $e_j > 0$ is not possible. To do so compute

$$\frac{\partial^2 E(V|\mathbf{y},\mathbf{e})}{\partial^2 e_i}|_{e_j = \hat{e} - e_i} = -\frac{\overline{q}_i}{E}\frac{\overline{q}_j}{E} E[V(y_i r_i^H + y_j r_j^H + \sum_{k \neq i,j} \tilde{r}_k y_k)] + \frac{\overline{q}_i}{E}\frac{\overline{q}_j}{E} E[V(y_i r_i^H + y_j \rho_j + \sum_{k \neq i,j} \tilde{r}_k y_k)] + \frac{\overline{q}_i}{E}\frac{\overline{q}_j}{E} E[V(y_i \rho_i + y_j r_j^H + \sum_{k \neq i,j} \tilde{r}_k y_k)] - \frac{\overline{q}_i}{E}\frac{\overline{q}_j}{E} E[V(y_i \rho_i + y_j \rho_j + \sum_{k \neq i,j} \tilde{r}_k y_k)].$$

We will show that if $V(c)$ is concave in c, then $\frac{\partial^2 E(V|\mathbf{y},\mathbf{e})}{\partial^2 e_i}|_{e_j = \hat{e} - e_i} > 0$, which means that

Table 1: Lotteries $\tilde{x}_L$ and $\tilde{x}_R$.

return on assets i, j	probability in $\tilde{x}_L$	probability in $\tilde{x}_R$
$r_i^H y_i + r_j^H y_j$	0	$\frac{1}{2}$
$r_i^H y_i + r_j^M y_j$	$\frac{1}{2}t_j$	0
$r_i^H y_i + r_j^L y_j$	$\frac{1}{2}(1-t_j)$	0
$r_i^M y_i + r_j^H y_j$	$\frac{1}{2}t_i$	0
$r_i^L y_i + r_j^H y_j$	$\frac{1}{2}(1-t_i)$	0
$r_i^M y_i + r_j^M y_j$	0	$\frac{1}{2}t_i t_j$
$r_i^M y_i + r_j^L y_j$	0	$\frac{1}{2}t_i(1-t_j)$
$r_i^L y_i + r_j^M y_j$	0	$\frac{1}{2}(1-t_i)t_j$
$r_i^L y_i + r_j^L y_j$	0	$\frac{1}{2}(1-t_i)(1-t_j)$

at the optimum either e_i^*, or e_j^* should be zero. First, note that $\frac{\partial^2 E(V|\mathbf{y},\mathbf{e})}{\partial^2 e_i}|_{e_j=\hat{e}-e_i} \geq 0$ is equivalent to

$$\begin{aligned}\frac{1}{2}E[V(y_i r_i^H + y_j\rho_j + \sum_{k\neq i,j}\tilde{r}_k y_k)] + \frac{1}{2}E[V(y_i\rho_i + y_j r_j^H + \sum_{k\neq i,j}\tilde{r}_k y_k)] \geq \\ \frac{1}{2}E[V(y_i r_i^H + y_j r_j^H + \sum_{k\neq i,j}\tilde{r}_k y_k)] + \frac{1}{2}E[V(y_i\rho_i + y_j\rho_j + \sum_{k\neq i,j}\tilde{r}_k y_k)].\end{aligned} \tag{26}$$

Recall that ρ_i is a binary random variable: $\Pr\{\rho_i = r_i^M\} = t_i$ and $\Pr\{\rho_i = r_i^L\} = 1 - t_i$ for all $i \in M$. Introduce random variable $\tilde{z} = \sum\limits_{k\neq i,j} \tilde{r}_k y_k$ which is independent of the returns on assets i, j. Note that the right-hand side of (26) corresponds to an expected utility from a compound lottery $\tilde{x}_R + \tilde{z}$, where $\tilde{x}_R$ represents the random return of assets i, j in this lottery. The left-hand side of (26) corresponds to an expected utility from a compound lottery $\tilde{x}_L + \tilde{z}$ where $\tilde{x}_L$ represents the random return of assets i, j in this lottery. Below we will show that $\tilde{x}_R$ is a mean-preserving spread of $\tilde{x}_L$.

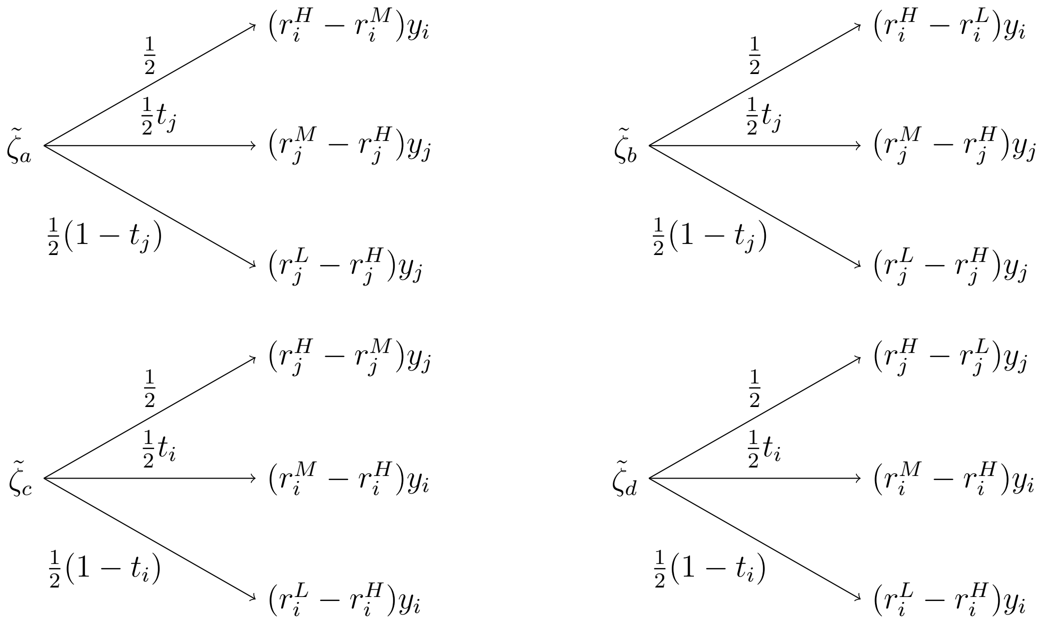
First, note that returns on assets $k \neq i, j$ do not matter for the comparison. Now, consider assets i and j. The table below describes possible realizations of monetary returns from assets i and j in lotteries $\tilde{x}_L$ and $\tilde{x}_R$ with corresponding probabilities.

It is easy to verify that both lotteries have the same expected monetary return. One can find a random variable $\tilde{\zeta}$ with zero mean such that $\tilde{x}_R = \tilde{x}_L + \tilde{\zeta}$, i.e. the right-hand side (RHS) lottery is a mean-preserving spread of the left-hand side (LHS) lottery. To see this, construct a compound lottery $\tilde{\zeta}$, that for each of the four final nodes of lottery $\tilde{x}_L$ specifies lotteries $\tilde{\zeta}_a$, $\tilde{\zeta}_b$, $\tilde{\zeta}_c$, $\tilde{\zeta}_d$ in the following manner:

$$\tilde{\zeta} = \begin{cases} \tilde{\zeta}_a & \text{if } \tilde{x}_L = r_i^M y_i + r_j^H y_j \text{ (probability } \frac{1}{2}t_i\text{)}, \\ \tilde{\zeta}_b & \text{if } \tilde{x}_L = r_i^L y_i + r_j^H y_j \text{ (probability } \frac{1}{2}(1-t_i)\text{)}, \\ \tilde{\zeta}_c & \text{if } \tilde{x}_L = r_i^H y_i + r_j^M y_j \text{ (probability } \frac{1}{2}t_j\text{)}, \\ \tilde{\zeta}_d & \text{if } \tilde{x}_L = r_i^H y_i + r_j^L y_j \text{ (probability } \frac{1}{2}(1-t_j)\text{)}. \end{cases}$$

Each lottery $\tilde{\zeta}_a$, $\tilde{\zeta}_b$, $\tilde{\zeta}_c$, $\tilde{\zeta}_d$ is played in the node which is reached with the corresponding probability in the LHS lottery described in the Table 1. To complete the proof, we need to find lotteries $\tilde{\zeta}_a$, $\tilde{\zeta}_b$, $\tilde{\zeta}_c$, $\tilde{\zeta}_d$ that map outcomes of the LHS lottery into outcomes of the RHS lottery; this can be done with the following lotteries.

Figure 6: Description of lotteries.



One can substitute and verify that $\tilde{x}_R = \tilde{x}_L + \tilde{\zeta}$. Since $r_k^H = \frac{1}{\pi_k} > 1 - \gamma_k = r_k^L$ and

$y_k > 0$ for $k = i, j$, lottery $\tilde{\zeta}$ is not degenerate. Its mean is zero:

$$\begin{aligned}
E[\tilde{\zeta}] &= \frac{1}{4}t_i[(r_i^H - r_i^M)y_i + t_j(r_j^M - r_j^H)y_j + (1-t_j)(r_j^L - r_j^H)y_j] \\
&+\frac{1}{4}(1-t_i)[(r_i^H - r_i^L)y_i + t_j(r_j^M - r_j^H)y_j + (1-t_j)(r_j^L - r_j^H)y_j] \\
&+\frac{1}{4}t_j[(r_j^H - r_j^M)y_j + t_i(r_i^M - r_i^H)y_i + (1-t_i)(r_i^L - r_i^H)y_i] \\
&+\frac{1}{4}(1-t_j)[(r_j^H - r_j^L)y_j + t_i(r_i^M - r_i^H)y_i + (1-t_i)(r_i^L - r_i^H)y_i] \\
&= \frac{1}{4}[t_i(r_i^H - r_i^M)y_i + (1-t_i)(r_i^H - r_i^L)y_i + t_j(r_j^M - r_j^H)y_j + (1-t_j)(r_j^L - r_j^H)y_j] \\
&+\frac{1}{4}[t_j(r_j^H - r_j^M)y_j + (1-t_j)(r_j^H - r_j^L)y_j + t_i(r_i^M - r_i^H)y_i + (1-t_i)(r_i^L - r_i^H)y_i] \\
&= \frac{1}{4}[r_i^H y_i - t_i r_i^M y_i - (1-t_i)r_i^L y_i - r_j^H y_j + t_j r_j^M y_j + (1-t_j)r_j^L y_j] \\
&+\frac{1}{4}[r_j^H y_j - t_j r_j^M y_j - (1-t_j)r_j^L y_j - r_i^H y_i + t_i r_i^M y_i + (1-t_i)r_i^L y_i] = 0.
\end{aligned}$$

Now consider separately the two cases of a risk-averse and a risk-neutral arbitrageur.

- If the arbitrageur is risk-averse, i.e. $V(c)$ is concave in c, then $\frac{\partial^2 E(V|\mathbf{y},\mathbf{e})}{\partial^2 e_i}|_{e_j=\hat{e}-e_i} > 0$. In this case the arbitrageur will never choose $e_i > 0$ and $e_j = \hat{e} - e_i > 0$, because setting $e_i = 0$ or $e_j = 0$ would increase the payoff. This implies that $e_i^* > 0$ and $e_j^* > 0$ cannot be optimal. In other words, $e_i^* = \hat{e} > 0$ for some $i \in M$ implies $e_j^* = 0$ for any $j \neq i$: only one asset is advertised by a risk-averse arbitrageur.

- If the arbitrageur is risk-neutral, i.e. $V(c) = c$, and the arbitrageur advertises both assets $e_i > 0$, $e_j > 0$, it must be the case that he invests in both assets $y_i > 0$, $y_j > 0$. It follows that both assets have the same expected return. Given that $q_i < 1$, $q_j < 1$ there is a profitable deviation for the arbitrageur. He can choose $e_i' = e_i + e_j$, $e_j' = 0$, $y_i' = y_i + y_j$, $y_j' = 0$ and benefit, because the return on asset i would increase due to extra advertising and, so the overall return on his investment would increase. Thus also a risk-neutral arbitrageur advertises only one asset.

It follows that in the relaxed problem the arbitrageur chooses $q_i^* = \overline{q}_i$ for some $i \in M$ and $q_j^* = 0$ for $j \neq i$.

ii) This optimal choice is also feasible in the initial problem, where $\mathbf{q} \in Q$ must satisfy (3). Hence the solution to the initial problem must have $q_i^* = \overline{q}_i$, $e_i^* = E$ for some $i \in M$ and $q_j^* = 0$, $e_j^* = 0$ for $j \neq i$. QED.

Proof of Lemma 1. When advertising is not possible, the arbitrageur's portfolio choice

$\mathbf{y}$ must satisfy his resource constraint $\sum_i y_i = w$ and maximize

$$t^2 E[V(y_k r^M + y_i r^M + \sum_{j \neq i,k} y_j \rho_j)] + t(1-t)E[V(y_k r^M + y_i r^L + \sum_{j \neq i,k} y_j \rho_j)]+ \\ (1-t)tE[V(y_k r^L + y_i r^M + \sum_{j \neq i,k} y_j \rho_j)] + (1-t)^2 E[V(y_k r^L + y_i r^L + \sum_{j \neq i,k} y_j \rho_j)] \tag{27}$$

The arbitrageur is risk-averse, so his objective is strictly concave in $\mathbf{y}$. The set of possible values is compact: $\sum_i y_i = w$, $y_i \geq 0$. Thus there exists an optimal portfolio and it is unique. Take asset k with $y_k^* \geq 0$ and fix $\bar{y} = y_i^* + y_k^*$, and y_j^* for $j \neq k, i$. Maximize (27) subject to $y_i = \bar{y} - y_k$ and y_j^* for $j \neq i, k$. The solution to this problem should deliver $y_k = y_k^*$ and $y_i^* = \bar{y} - y_k^*$. The first order condition is

$$t^2 E[V'(y_k r^M + (\bar{y} - y_k) r^M + \sum_{j \neq i,k} y_j^* \rho_j)](r^M - r^M)+ \\ t(1-t)E[V'(y_k r^M + (\bar{y} - y_k) r^L + \sum_{j \neq i,k} y_j^* \rho_j)](r^M - r^L)+ \\ (1-t)tE[V'(y_k r^L + (\bar{y} - y_k) r^M + \sum_{j \neq i,k} y_j^* \rho_j)](r^L - r^M)+ \\ (1-t)^2 E[V'(y_k r^L + (\bar{y} - y_k) r^L + \sum_{j \neq i,k} y_j^* \rho_j)](r^L - r^L) = 0. \tag{28}$$

As the first term and the last term of the left-hand side of (28) are zero, equation (28) becomes: $E[V'(y_k r^M + (\bar{y} - y_k) r^L + \sum_{j \neq i,k} y_j^* \rho_j)] = E[V'(y_k r^L + (\bar{y} - y_k) r^M + \sum_{j \neq i,k} y_j^* \rho_j)]$, which implies $y_k^* = y_i^* = \bar{y}/2$. One can verify that corner solutions $y_k = 0$, $y_k = \bar{y}$ do not satisfy the necessary condition because V is concave. A similar argument for any pair of other assets i and $j \neq i$ would imply $y_j^* = y_i^*$. As the number of assets in M is M, we get $y_i^* = w/M$. QED.

Proof of Proposition 2. First, the arbitrageur must advertise the asset with the greatest advertisability $i = \arg\max_{k \in M} \overline{q}_k$. Suppose otherwise $e_i = 0$ and $e_j = E$ for some $j \neq i$. Denote corresponding investments y_i and $y_j = y > 0$. This is not optimal, because the arbitrageur can increase his utility by switching around both advertising effort and investment levels between the two assets, namely, by setting $y_i' = y_j = y$, $y_j' = y_i$, $e_i' = e_j = E$ and $e_j' = e_i = 0$. Indeed, investments y_j' and y_i deliver identical returns. However, investment y_i' dominates investment y_j in terms of first order stochastic dominance, as the table below illustrates:

Table 2: Investments y'_i and y_j.

return on investment	probability for $y'_i = y$, $e'_i = E$	probability for $y_j = y$, $e_j = E$.
$r^H y$	$\overline{q}_i$	$\overline{q}_j$
$r^M y$	$(1-\overline{q}_i)t$	$(1-\overline{q}_j)t$
$r^L y$	$(1-\overline{q}_i)(1-t)$	$(1-\overline{q}_j)(1-t)$

Since $i = \underset{k \in M}{\arg\max}\, \overline{q}_k$, it is optimal to set $e_i = E$.

Second, it is straightforward to show that the arbitrageur invests equal amounts in the assets that he does not advertise. The argument is the same as in the proof of Lemma 1.

To prove that $y_i > y_j$, $j \neq i$, rewrite the arbitrageur's expected utility as follows:

$$
\begin{aligned}
E[V|\mathbf{y},\mathbf{e}] &= \overline{q}_i t E[V(y_i r^H + y_j r^M + \sum_{k \neq i,j} y_k \rho_k)] + \\
&\overline{q}_i (1-t) E[V(y_i r^H + y_j r^L + \sum_{k \neq i,j} y_k \rho_k)] + \\
&(1-\overline{q}_i) t^2 E[V(y_i r^M + y_j r^M + \sum_{k \neq i,j} y_k \rho_k)] + \\
&(1-\overline{q}_i) t(1-t) E[V(y_i r^M + y_j r^L + \sum_{k \neq i,j} y_k \rho_k)] + \\
&(1-\overline{q}_i)(1-t) t E[V(y_i r^L + y_j r^M + \sum_{k \neq i,j} y_k \rho_k)] + \\
&(1-\overline{q}_i)(1-t)^2 E[V(y_i r^L + y_j r^L + \sum_{k \neq i,j} y_k \rho_k)].
\end{aligned}
\tag{29}
$$

As before, we fix all optimal y_k^*, $k \neq j, i$ and set $\bar{y} = y_i^* + y_j^* > 0$. Consider then optimization of (29) over y_i given the constraint $y_j = \bar{y} - y_i$. The first order necessary condition with respect to y_i is:

$$
\begin{aligned}
&\overline{q}_i t E[V'(y_i r^H + y_j r^M + \sum_{k \neq i,j} y_k \rho_k)](r^H - r^M) + \\
&\overline{q}_i (1-t) E[V'(y_i r^H + y_j r^L + \sum_{k \neq i,j} y_k \rho_k)](r^H - r^L) + \\
&(1-\overline{q}_i) t^2 E[V'(y_i r^M + y_j r^M + \sum_{k \neq i,j} y_k \rho_k)](r^M - r^M) + \\
&(1-\overline{q}_i) t(1-t) E[V'(y_i r^M + y_j r^L + \sum_{k \neq i,j} y_k \rho_k)](r^M - r^L) + \\
&(1-\overline{q}_i)(1-t) t E[V'(y_i r^L + y_j r^M + \sum_{k \neq i,j} y_k \rho_k)](r^L - r^M) + \\
&(1-\overline{q}_i)(1-t)^2 E[V'(y_i r^L + y_j r^L + \sum_{k \neq i,j} y_k \rho_k)](r^L - r^L) = 0.
\end{aligned}
\tag{30}
$$

This reduces to

$$\overline{q}_i t E[V'(y_i r^H + y_j r^M + \sum_{k \neq i,j} y_k \rho_k)](r^H - r^M) +$$
$$\overline{q}_i (1-t) E[V'(y_i r^H + y_j r^L + \sum_{k \neq i,j} y_k \rho_k)](r^H - r^L) =$$
$$(r^M - r^L)(1 - \overline{q}_i)(1-t) t E[V'(y_i r^L + y_j r^M + \sum_{k \neq i,j} y_k \rho_k) - V'(y_i r^M + y_j r^L + \sum_{k \neq i,j} y_k \rho_k)].$$

The LHS is positive. Given that V is concave the RHS is positive if and only if $y_i r^L + y_j r^M < y_i r^M + y_j r^L$, which implies $y_i > y_j$. QED.

Proof of Proposition 5. If advertising is successful at $t = 1$ and correctly predicts the asset's return at $t = 2$, then using (21) we express the price reaction at $t = 4$ to successful advertising as $\Delta(1,1) = (1 - \pi_j)\frac{\mu}{\mu + (1-\mu)\pi_i}$. If advertising fails at $t = 1$, then again using (21) we can express the price reaction at $t = 4$ to successful advertising as $\Delta(\emptyset) = (1 - \pi_j)\mu < \Delta(1,1)$. Finally, if advertising is successful at $t = 1$ but it is contradicted by the asset's return at $t = 2$, then $\mu_2(1,0) = 0$, so that according to (21) the price at $t = 4$ will not react to advertising: $\Delta(1,0) = 0$. 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.

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.

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.

Ljungqvist, A. and W. Qian (2014): "How Binding Are Limits to Arbitrage?" *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 UnderDiversification," *Review of Economic Studies*, 77, 779–805.

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