EIEF WORKING PAPER SERIES

IEF

EINAUDI INSTITUTE FOR ECONOMICS AND FINANCE

EIEF Working Paper 01/14
April 2014

Advertising Arbitrage

by

Sergei Kovbasyuk
(EIEF)
Marco Pagano
(University of Naples Federico II, CSEF & EIEF)

Advertising Arbitrage*

Sergei Kovbasyuk † Marco Pagano ‡

July 31, 2014

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, prices react more strongly to the ads of arbitrageurs with a successful track record, and reputation-building induces high-skill arbitrageurs to advertise more than others.

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

JEL classification: G11, G14, G2, D84.

*We are grateful to Bruno Biais, Thierry Foucault, Mikhail Golosov, Hugo Hopenhayn, Tullio Jappelli, 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 EIEF, CSEF and to 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. But, once the arbitrageur has invested in an asset, his advertising effort will be greater if future public information about it is imprecise, to compensate for the poor quality of public information.

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. And in the dynamic version of our model, 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 fades. Reputation-building also motivates high-skill arbitrageurs to advertise more than low-skill ones, as they anticipate that they will be more likely to reap large gains in the future. 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

[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.

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 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.

[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.

Section 5 allows for strategic interactions among arbitrageurs. Section 6 investigates the way in which arbitrageurs' reputation affects the effectiveness of their advertising, first in a static setting and then in a dynamic one where arbitrageurs build their reputation over time. The last section summarizes and discusses our predictions.

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.

At $t = 1$ a noisy public signal $s_i \in \{0, 1\}$ about θ_i, $i \in N$ becomes available. 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 $t = 0$, the arbitrageur privately learns θ_i for a finite subset of assets $i \in M$ and decides in which assets to take positions.[4] The arbitrageur can try to communicate θ_i about any asset $i \in M$ to investors by exerting "advertising effort" $e_i \geq 0$. This effort captures all expenses born by the arbitrageur to prove and expose the mispricing, including the collection of hard evidence and its dissemination. 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.

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 learn the true realization of θ_i at $t = 1$ only if advertising is effective, which happens with probability $q_i = \min[a_i e_i, 1]$, $a_i \in (0, 1]$ for any $i \in M$.*

With complementary probability, advertising fails and investors learn the true θ_i only at $t = 2$. Parameter a_i captures the extent to which information about asset i is "advertisable", and thus represents the various factors that facilitate the collection of evidence about the asset and its communication to investors. For instance, a_i may be high when

[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.

investors are very receptive to information about asset i, either 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.

At $t = 0$ 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. The timeline is as follows (see also Figure 1).

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 $t = 1$ for each asset $i \in M$, the public signal s_i is realized. With probability $q_i = \min[a_i e_i, 1]$ the arbitrageur's advertising is effective and investors learn θ_i; with complementary probability, investors rely on s_i. Each asset i can be traded at p_1^i, so that the arbitrageur's monetary payoff is $c = \sum\limits_i x_i p_1^i$. Finally, at $t = 2$ all assets' final returns θ_i, $i \in M$ are realized.

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

0	1	2
Asset price: p_0^i	Public signal s_i	θ_i is realized
Arbitrageur's position: x_i	Investors learn θ_i with probability q_i	
Arbitrageur's effort: e_i	Asset price: p_1^i	
	Arbitrageur liquidates	

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.

The arbitrageur's utility $V(c, e)$ at $t = 1$ is a function of his monetary payoff $c = \sum\limits_i x_i p_1^i$ at that time and his total advertising effort $e = \sum\limits_i e_i \geq 0$. The utility function is increasing in the monetary payoff, decreasing in advertising effort, and not convex: $V_c > 0$, $V_{cc} \leq 0$, $V_e < 0$ for $e > 0$, $V_e(c, 0) = 0$, $V_{ee} \leq 0$. The cost of advertising is not affected by the monetary payoff $V_{ec} = 0$.

Assumption 2. *The arbitrageur has limited resources $w > 0$ at $t = 0$.*

At $t = 0$ the arbitrageur can allocate resources w 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 (1) also imposes a constraint on the arbitrageur's short positions, because in practice both long and short positions require some collateral.

We assume that the arbitrageur's 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. *Arbitrageur's trades do not affect prices.*

For brevity, and without loss of generality, we consider the case of undervalued assets:

Assumption 4. *All assets in M are undervalued $\theta_i = 1$, $i \in M$.*

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

We posit a limit on arbitrageurs' interest in advertising an asset.

Assumption 5. *Perfect advertising is prohibitively costly: $V(\frac{w}{\underline{\pi}}, 1) - V(0,1) < |V_e(\frac{w}{\underline{\pi}}, 1)|$.*

This assumption is equivalent to the following condition: $\frac{\partial}{\partial e_i}[q_i V(\frac{w}{\underline{\pi}}, e_i) + (1-q_i)V(0, e_i)] < 0$ for $e_i = 1$, which ensures that even if the arbitrageur invested all his wealth w in the most underpriced asset ($p_i^0 = \underline{\pi}$), and this asset was the easiest to advertise ($a_i = 1$), he still would not choose an advertising level $e_i = 1$ such that $q_i = 1$, i.e. investors learn θ_i for sure. In other words, the marginal cost of advertising effort $e_i = 1$ is sufficiently high. This natural assumption simplifies the analysis, as we can take it for granted that $q_i < 1$ for any $i \in M$.

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

$$\tilde{r}_i = \begin{cases} r_i^H & \text{with probability } q_i \\ r_i^M & \text{with probability } (1 - q_i)t_i \\ r_i^L & \text{with probability } (1 - q_i)(1 - t_i) \end{cases}, i \in N. \quad (2)$$

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 (1) and (2) 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 $\rho_i \in \{r_M, r_L\}$, with $\Pr(\rho = r_M) = t_i$. 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:

$$\begin{aligned} &E[V|\mathbf{y}, \mathbf{e}] = q_i q_j E[V(y_i r_i^H + y_j r_j^H + \tilde{r}_{-ij}, e)] + q_i(1 - q_j)E[V(y_i r_i^H + y_j \rho_j + \tilde{r}_{-ij}, e)] + \\ &(1 - q_i)q_j E[V(y_i \rho_i + y_j r_j^H + \tilde{r}_{-ij}, e)] + (1 - q_i)(1 - q_j)E[V(y_i \rho_i + y_j \rho_j + \tilde{r}_{-ij}, e)]. \end{aligned} \quad (3)$$

The portfolio choice and advertising decisions solve:

$$\max_{\{\mathbf{y} \geq 0, \mathbf{e} \geq 0\}} E[V|\mathbf{y}, \mathbf{e}], \ s.t. \ \sum_i y_i \leq w, \ q_i = \min[a_i e_i, 1], \ \forall i \in M. \quad (4)$$

We start to solve the arbitrageur's problem by characterizing his advertising decisions.

Lemma 1. *In any solution of the arbitrageur's problem, advertising never succeeds with certainty:* $q_i < 1$ *for all* i.

All proofs are in the appendix. Recall that by Assumption 5 the marginal cost of advertising is high enough that the arbitrageur never advertises an asset so much that investors certainly learn θ_i, i.e. so much that $q_i = 1$.

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

The proof 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. The detailed proof is in the appendix. We illustrate the intuition with a simple symmetric example with two identical assets and an uninformative public signal.

Example with two assets. M contains two identical assets $i = 1, 2$ such that $\gamma_1 = \gamma_2 = 0$, $r_1^L = r_2^L = 1$, $r_1^H = r_1^M = r_2^H = r_2^M = r > 1$, $a_1 = a_2 = 1$. 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.

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, $q_j = 0$ for $j \neq i$ and the expression for the arbitrageur's utility (3) can be written as

$$E[V|\mathbf{y}, e_{-i} = 0] = q_i E[V(y_i r_i^H + \sum_{j \neq i} \rho_j y_j, e_i)] + (1 - q_i) E[V(\sum_{j \in M} \rho_j y_j, e_i)]. \quad (5)$$

The arbitrageur's optimization problem (4) can be solved as follows. For each $i \in M$, find $\mathbf{e}(i)$ and $\mathbf{y}(i)$ that maximize (5) subject to $\sum_i y_i \leq w$ and $q_i = a_i e_i$. By Lemma 1 $q_i < 1$, so that we can consider $e_i \in [0, 1/a_i]$ without loss of generality. For any given $e_i \in [0, 1/a_i]$, $q_i = a_i e_i$ is fixed and one can find a portfolio $\mathbf{y}(e_i)$ that maximizes (5) subject to $\sum_i y_i \leq w$. For each e_i, denote by $E[V|e_i]$ the corresponding maximal value. The function $E[V|e_i]$ is bounded for $e_i \in [0, 1/a_i]$ and therefore achieves a maximum for some e_i^*. Denote the maximal value $E[V]_i$, which may be achieved by multiple levels of e_i. By advertising asset $i \in M$, the arbitrageur can get at most $E[V]_i$. At the optimum, he advertises asset $i^* \in \arg\max_j E[V]_j$, and there may be multiple assets that deliver the same maximal payoff. The level of advertising is $e_{i^*}^*$ and the portfolio choice is $\mathbf{y}(e_{i^*}^*)$.

As the above argument illustrates, once the arbitrageur has chosen the asset i^* and his advertising effort $e^*_{i^*}$, portfolio choice becomes a standard diversification problem. The only difference is that the likelihood of a high return on investment in 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 make 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 6. *Assets in M differ only in terms of advertisability: $\gamma_i = \gamma$ and $\pi_i = \pi$ for all $i \in M$ and $a_i \neq a_j$ 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 2. *When advertising is not possible, the arbitrageur is risk-averse, and Assumption 6 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. When advertising is possible, this is not the case.

Proposition 2. *When advertising is possible, the arbitrageur is risk-averse, and Assumption 6 holds, the arbitrageur advertises the most advertisable asset and invests more in it than in any other asset: for $i = \arg\max_{j \in M} a_j$ 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 a highest expected return for a given level of 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 c and that his effort cost function is quadratic.

Assumption 7. $V(c, e) = c - e^2/2$.

However, we drop Assumption 6 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 ($a_i \neq a_j$ 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 $q_k = a_k e_k$ for $e_k \leq 1/a_k$, advertising effort should maximize the expected payoff:

$$\max_{e_k \in [0, 1/a_k]} a_k e_k r_k^H w + (1 - a_k e_k)[t_k r_k^M + (1 - t_k) r_k^L] w - e_k^2/2. \tag{6}$$

From the first order condition, the optimal advertising effort is[5]

$$e_k^* = a_k (1 - \gamma_k^2) \frac{1 - \pi_k}{\pi_k} w. \tag{7}$$

Remark 1. *The optimal advertising effort increases with the asset's "advertisability" (a_k) and mispricing ($1/\pi_k$), and decreases with the precision of the public signal (γ_k).*

Intuitively, a unit of advertising effort is more productive for more "advertisable" assets and more profitable for those that are more mispriced. Noisier public information (lower

[5] Note that, by Assumption 5, $w/\underline{\pi} < 1$ and the solution is interior: $a_k e_k^* < 1$.

γ_k) induces arbitrageurs to advertise more aggressively and speed up price discovery: more advertising substitutes for poorer public information.[6]

The optimal choice of effort in equation (7) is conditional on the arbitrageur picking asset k. What guides the choice of asset k is the expected payoff from investing in it and advertising it:

$$E[V|\pi_k, \gamma_k, a_k] = w\left[1 + \gamma_k^2\left(\frac{1}{\pi_k} - 1\right)\right] + \frac{w^2 a_k^2}{2}\left(1 - \gamma_k^2\right)^2\left(\frac{1}{\pi_k} - 1\right)^2. \tag{8}$$

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

The proof is straightforward. All three characteristics (potential return $\frac{1}{\pi_k}$, advertisability a_k, and quality of public information γ_k) are desirable from the arbitrageur's point of view:

$$\begin{aligned}
\frac{\partial E[V|\pi, \gamma, a]}{\partial(1/\pi)} &= w\gamma^2 + a^2 w^2 (1-\gamma^2)^2\left(\frac{1}{\pi} - 1\right) > 0, \\
\frac{\partial E[V|\pi, \gamma, a]}{\partial a} &= w^2 a (1-\gamma^2)^2\left(\frac{1}{\pi} - 1\right)^2 > 0, \\
\frac{\partial E[V|\pi, \gamma, a]}{\partial \gamma} &= 2\gamma w\left(\frac{1}{\pi} - 1\right)\left[1 - wa^2(1-\gamma^2)\left(\frac{1}{\pi} - 1\right)\right] > 0.
\end{aligned} \tag{9}$$

The first two inequalities are obvious, and the last follows from Assumption 5, which guarantees $w/\pi < 1$. 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 a_i and low potential return $1/\pi_i$, and one with low advertisability a_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

[6] Interestingly, one of the arbitrageurs that engage in aggressive advertising has chosen the name "Muddy Waters Research".

information.

5 Multiple arbitrageurs

When several arbitrageurs acquire private information about different assets independently, each of them behaves as described in 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. The advertising effort of each contributes to the success of advertising: we assume that for any asset $i \in M$ advertised by several arbitragers $e_i^l \geq 0$, $l \in M$, the probability of investors learning the true θ_i at $t = 1$ is $q_i = a_i \sum_l e_i^l$. As before, we want to avoid perfect advertising $q_i = 1$, so we modify Assumption 5 to adapt it to the presence of multiple arbitrageurs and assume $Lw < \underline{\pi}$.

The possible realizations of the return on investment in asset i are characterized by equation (2), 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 can be advertised in equilibrium. Third, we provide an example that shows that the equilibrium can be "inefficient": the arbitrageurs would be better off if they all invested in a different asset and advertised it.

Lemma 3. *In equilibrium all L arbitrageurs invest in the same asset.*

To prove this, it suffices to show that 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. If an arbitrageur who invests in asset j deviated, by investing 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 show that in equilibrium, if arbitrageurs invest in asset j and advertise it, no arbitrageur wants to deviate. If the arbitrageur deviates, he chooses an asset different from j that maximizes his expected payoff in autarky. Denote this asset by $h_j = \underset{k \in M \backslash j}{\arg\max}$:

$E[V|\pi_k, \gamma_k, a_k]$. The corresponding expected payoff is

$$V_{-j}^a = w\left[1+\gamma_{h_j}^2\left(\frac{1}{\pi_{h_j}}-1\right)\right]+\frac{w^2a_{h_j}^2}{2}(1-\gamma_{h_j}^2)^2\left(\frac{1}{\pi_{h_j}}-1\right)^2. \tag{10}$$

If all arbitrageurs invest in asset j, each arbitrageur $l \in L$ chooses his advertising effort in order to maximize his expected payoff:

$$\max_{e_j^l \in [0,1/a_j]} q_j r_j^H w + (1-q_j)[t_j r_j^M + (1-t_j) r_j^L]w - (e_j^l)^2/2,$$

where $q_j = a_j(e_j^l + \sum_{m \neq l} e_j^m)$. As above, if advertising succeeds, the return at $t = 1$ is $r_j^H = 1/\pi_j$; if it fails and the public signal is $s_i = 1$, the return is $r_j^M = 1 - \gamma_j + \frac{\gamma_j}{\pi_j}$; and if it fails and $s_i = 0$, the return is $r_j^L = 1 - \gamma_j$. From the first order condition, every arbitrageur chooses the optimal effort $e_j = a_j(1-\gamma_j^2)\frac{1-\pi_j}{\pi_j}w$, so that the probability of successful advertising is $q_j = La_j^2(1-\gamma_j^2)\frac{1-\pi_j}{\pi_j}w$, where the assumption $Lw < \underline{\pi}$ guarantees $q_j < 1$. Substituting for advertising efforts, we obtain each arbitrageur's expected payoff if all arbitrageurs invest in asset j and advertise it:

$$V_j(L) = w\left[1+\gamma_j^2\left(\frac{1}{\pi_j}-1\right)\right]+\left(L-\frac{1}{2}\right)w^2a_j^2(1-\gamma_j^2)^2\left(\frac{1}{\pi_j}-1\right)^2. \tag{11}$$

It is easy to see that condition $V_j(L) \geq V_{-j}^a$ ensures that every arbitrageur prefers to invest in the asset that is already advertised by the other $L-1$ arbitrageurs.

Proposition 4. *An equilibrium in which all arbitrageurs invest in asset j and advertise it exists if and only if $V_j(L) \geq V_{-j}^a$.*

The proof is immediate. It is also easy to see that multiple equilibria may be possible: the condition $V_j(L) \geq V_{-j}^a$ can hold for several $j \in M$. This multiplicity arises from the strategic complementarity between arbitrageurs: each has the incentive to "piggyback" on the advertising of others. Since an asset that is 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 seem to resemble the herding induced by information cascades, but in fact it is quite different: in this model, the fact that all arbitrageurs pick the same 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.

The multiplicity of equilibria may become extreme if many arbitrageurs have the same information about mispriced assets:

Corollary 1. *Any asset $j \in M$ for which arbitrageurs have information can be advertised in equilibrium by all of them, if the number of arbitrageurs L exceeds a critical threshold $\underline{L}(j) < \infty$ and their individual resources are limited $w < \underline{\pi}/L$.*

Intuitively, when arbitrageurs are most numerous, the strategic complementarity between them is strongest: as a result, in equilibrium they may all concentrate their investment and advertising efforts on any asset $j \in M$, whether its price is far from the fundamental value or not, and whether it is easy to advertise or not. Hence, in equilibrium they may choose an asset that is only moderately mispriced ($\pi_j = \overline{\pi}$) and relatively difficult to advertise ($a_j \to 0$), even if there exists another asset $i \in M$, that is much more severely underpriced and much easier to advertise ($\pi_i = \underline{\pi} < \pi_j$, $a_i = 1 > a_j$ and $\gamma_i = \gamma_j$). Such an equilibrium is inefficient: investors would jointly prefer to coordinate on asset i rather than on asset j.

However, inefficiency does not require a large number of arbitrageurs: the following example shows that even with just two ($L = 2$) the equilibrium can be inefficient.

Example with two arbitrageurs and two assets. Take $L = 2$. Consider assets $i = 1, 2$ and assume $\pi_1 < \pi_2$, $\gamma_1 = \gamma_2 = 0$, $a_1 = a_2 = 1$. Notice that asset 1, other things being equal, delivers a higher potential return than asset 2.

Suppose both arbitrageurs invest in asset 1 in equilibrium, then from (10) each gets $V_1(2) = w + \frac{3}{2}w^2\left(\frac{1}{\pi_1} - 1\right)^2$. If one deviates, invests in asset 1 and advertises it, then by (11) he gets $V^a_{-1} = w + \frac{1}{2}w^2\left(\frac{1}{\pi_2} - 1\right)^2$. Note that $\pi_1 < \pi_2$ implies $V_1(2) \geq V^a_{-1}$ and such an equilibrium always exists, by Proposition 4.

Suppose both arbitrageurs invest in asset 2 in equilibrium, then $V_2(2) = w+\frac{3}{2}w^2\left(\frac{1}{\pi_2} - 1\right)^2$ and $V^a_{-2} = w+\frac{1}{2}w^2\left(\frac{1}{\pi_1} - 1\right)^2$. Such an equilibrium exists if and only if $V_2(2) \geq V^a_{-2}$, which is equivalent to $\frac{1}{\pi_1} - 1 \leq \sqrt{3}\left(\frac{1}{\pi_2} - 1\right)$. If this condition holds, then both equilibria exist.

It is easy to see that the arbitrageurs prefer the equilibrium in which they both invest in asset 1 and advertise it, indeed $\pi_1 < \pi_2$ implies $V_1(2) > V_2(2)$. In this equilibrium they invest in asset 1 with the greatest mispricing $1/\pi_1$ and advertise it. Yet the other equilibrium is also possible: if both advertise asset 2, neither will want to deviate and advertise asset 1. The latter equilibrium is inefficient, because the arbitrageurs would benefit if they could coordinate on investment in asset 1 and advertise it.

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. Hence, this strategic complementarity 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 setup 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 Credibility

Until now we have posited an arbitrageur with perfect information about asset returns θ_i, $i \in M$. On this assumption, with successful advertising of asset i 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 the information depends on the credibility of the arbitrageur. For simplicity, we assume that the arbitrageur has private information on a single asset. Other investors assign the same probability to the arbitrageur having information about any particular asset in N. We further 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 $\tau = L$, 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 start with, in section 6.1 we study a static model that excludes reputation-building. Section 6.2 extends the analysis to a setting where investors update their beliefs about the arbitrageur's type based on his previous performance, thus allowing for reputation-building. Here the arbitrageur takes into account how investors' belief μ about his type evolves depending on the information that he advertises and on how well it matches the actual realization.

6.1 Static model

In the static case, the timeline is as in the basic model. 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 chooses advertising effort $e_\tau \in [0, 1/a_i]$. 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 $q = a_i e_\tau$; 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 realized returns.

We solve for the Perfect Bayesian Equilibrium of this game. In equilibrium, at $t = 0$ asset prices are determined by investors' prior beliefs, so that asset i trades at price $p_0^i = \pi_i$. At $t = 1$ the price is the expected asset value conditional on the signal and on the arbitrageur's credibility, i.e. the investors' posterior belief $\mu(\hat{s}_i)$ about his accuracy:

$$p_1^i = E_\mu[\theta_i|\hat{s}_i], \ i \in N. \tag{12}$$

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 $\mu(\hat{s}_i) = 0$ if they receive the arbitrageur's signal $\hat{s}_i = \hat{\theta}_i$, so that advertising has no effect on prices, and the arbitrageur does not find it optimal to advertise, irrespective of his type. There is also an equilibrium in which the arbitrageur advertises, his signal raises the price at $t = 1$ (so that $p_1^i > p_0^i$), and the arbitrageur, being risk-neutral, invests his entire wealth in asset i ($x = w/\pi_i$). Multiple equilibria imply that there may be situations in which arbitrageurs never get the chance to develop credibility because of general skepticism about their skill, and others in which their advertising has credibility. In what follows, we concentrate on the more interesting equilibrium, the one with advertising.

The equilibrium with advertising is fully characterized by the investors' beliefs $\mu(\hat{s}_i)$ and the arbitrageur's advertising effort e_τ. High-skill and low-skill arbitrageurs choose their efforts e_H^* and e_L^* optimally, given investors' beliefs, which in turn must be consistent with Bayes' rule.

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

$$\mu_1(F) = \frac{\mu(1 - a_i e_H^*)}{\mu(1 - a_i e_H^*) + (1 - \mu)(1 - a_i e_L^*)}, \tag{13}$$

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:

$$p_1^i(F) = E[\theta_i] = \pi_i. \tag{14}$$

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

$$\mu_1(S) = \frac{\mu a_i e_H^*}{\mu a_i e_H^* + (1 - \mu) a_i e_L^*}, \tag{15}$$

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 i at $t = 1$ is

$$p_1^i(S) = \mu_1(S) + (1 - \mu_1(S))\pi_i, \tag{16}$$

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_i.

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

$$\Delta \equiv p_1^i(S) - p_1^i(F) = \mu_1(S)(1 - \pi_i). \tag{17}$$

At $t = 0$ the arbitrageur takes the price reaction into account when he decides on advertising effort e. In expectation, asset i's price at $t = 1$ is $E[p_1^i|e] = a_i e p_1^i(S) + (1 - a_i e)p_1^i(F) = \pi + a_i e\Delta$, which can be expressed as $E[p_1^i|e] = \pi_i + a_i e \mu_1(S)(1 - \pi_i)$ using equations (14) and (16). By Assumption 7 the arbitrageur's expected payoff as of $t = 0$ is $E[p_1^i|e]x + (w - xp_0^i) - e^2/2$. Given that $p_0^i = \pi_i$ at $t = 0$, the arbitrageur's expected payoff is $a_i e\Delta x + w - e^2/2$, which is increasing in x if $e > 0$. Hence the type τ arbitrageur invests his entire wealth w in asset i ($x = w/\pi_i$) and chooses advertising effort so as to maximize:

$$\max_{e_\tau \in [0,1/a_i]} \frac{a_i e_\tau \Delta}{\pi_i} w + w - e_\tau^2/2, \ \tau \in \{L, H\}. \tag{18}$$

Problem (18) is convex, so that the first order condition delivers the optimal effort $e^* = wa_i\Delta/\pi_i = wa_i\mu_1(S)(1-\pi_i)/\pi_i$ for any τ. The result is highly intuitive. At $t = 1$ the investors' beliefs about the arbitrageur $\mu(S)$ and $\mu(F)$ are the same, regardless of the arbitrageur's actual skill τ. Hence, both types of arbitrageur have the same incentives, so that their advertising efforts are the same. This in turn implies that the arbitrageur's success or failure in advertising does not convey any information about his type, so that the investors' posterior belief coincides with their prior $\mu_1(S) = \mu = \mu_1(F)$. Hence, the price reaction to advertising Δ is fully determined by the investors' prior belief about the arbitrageur's skill μ: $\Delta^* = \mu(1-\pi_i)$. This proves:

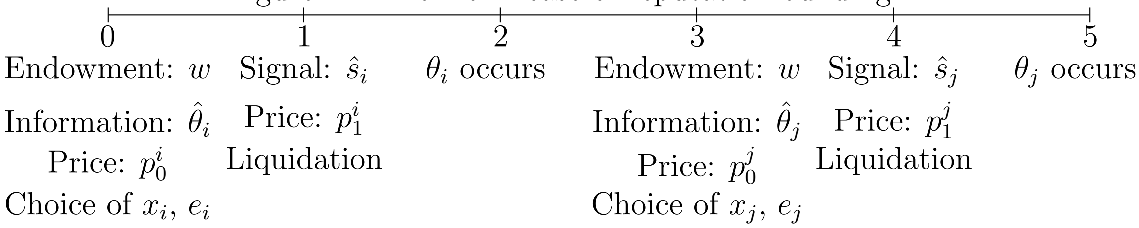
Proposition 5. *In equilibrium, the two types of arbitrageur exert the same level of effort. Equilibrium effort and price reaction to advertising increase with the arbitrageur's credibility μ and with the extent of mispricing $1-\pi_i$.*

The result that the arbitrageur exerts the same effort irrespective of his type is an artifact of the static model. In the dynamic model that we study in the next section high-skill arbitrageur exerts more advertising effort than the low-skill arbitrageur, and the investors' posterior beliefs about the arbitrageur's type depend on whether advertising is successful or not.

6.2 Reputation-building

Thus far we have assumed that after liquidating his position the arbitrageur does not reinvest again. This is reasonable in the basic setup where arbitrageurs are publicly known to possess reliable private information and there is no scope for reputation-building. In the current setup with uncertainty about arbitrageur's skill level, however, it is important to explore how our findings change if we introduce the possibility of repeated interaction and reputation-building. We propose a simple extension of the model: at the end of period $t = 2$, after actual asset returns are realized, the whole sequence of actions described at the beginning of section 6.1 is repeated. That is, at $t = 3$ a new set N of assets with uncertain returns appears. The arbitrageur has a new endowment w and observes a signal about an asset j in N. The whole timeline of actions and events then unfolds as before: the arbitrageur decides on new investment and on advertising. At $t = 4$ the advertising either succeeds or not, asset prices adjust, the arbitrageur liquidates and consumes. At $t = 5$ the new actual returns are realized.

Figure 2: Timeline in case of reputation building.



For simplicity, we assume that if advertising fails at $t = 1$ ($\hat{s}_i = \emptyset$), i.e. if investors ignore the arbitrageur's signal $\hat{\theta}_i$, the signal $\hat{\theta}_i$ that they missed cannot subsequently be retrieved.[7] 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. To simplify the notation, with no loss of generality we standardize the arbitrageur's wealth to be just enough to purchase one unit of the undervalued asset, and set the asset's advertisability at its maximal level:

Assumption 8. $w = \pi = \pi_i = \pi_j$ *and* $a_i = a_j = 1$.

Notice that the interaction from $t = 3$ onward is equivalent to the static model with reputation, with one difference: the investors' posterior belief about the arbitrageur's type μ_3 at $t = 3$ may depend on the past realization of asset i's return θ_i at $t = 2$. Since no new information arrives at $t = 3$, we have $\mu_3 = \mu_2$. The arbitrageur's problem at $t = 3$ is exactly the same as problem (18), simply replacing investors' prior belief μ with their posterior belief μ_2. As before, we concentrate on the equilibrium with advertising. Proposition 5, taken together with Assumption 8, implies that at $t = 3$ an arbitrageur with reputation μ_2 will exert effort $e = \mu_2(1 - \pi)$. Substituting this into the arbitrageur's payoff function, we obtain his expected utility at the beginning of period $t = 3$:

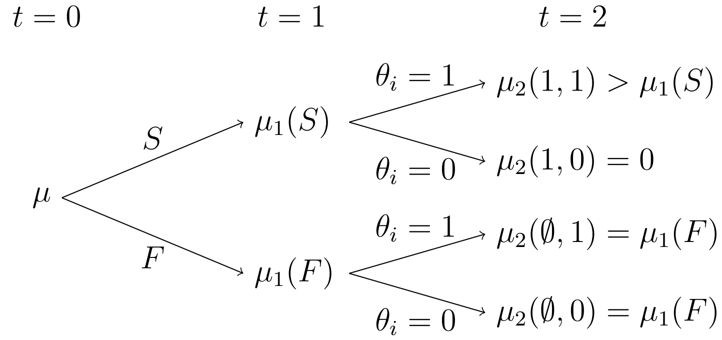
$$V'(\mu_2) = w + (\mu_2(1 - \pi))^2 . \tag{19}$$

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 continuation payoff V' at $t = 3$. We need to describe how investors update their beliefs about the arbitrageur when at $t = 1$ they learn whether advertising succeeded (S) or not (F), and when at $t = 2$ they

[7] 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.

observe the actual return θ_i. Figure 3 below summarizes the evolution of investors' beliefs.

Figure 3: Evolution of investors' beliefs



Posterior beliefs $\mu_1(F)$ and $\mu_1(S)$ at $t = 1$ are characterized by equations (13) and (15). To describe the posterior belief $\mu_2(\hat{s}_i, \theta_i)$ at $t = 2$, note first that if advertising fails (F), investors do not observe $\hat{\theta}_i$ $(\hat{s}_i = \emptyset)$ and cannot compare it with the actual realization of θ_i, so that in this case they do not update their beliefs at $t = 2$: $\mu_2(\emptyset, 1) = \mu_2(\emptyset, 0) = \mu_1(F)$. Recall that at $t = 0$ the arbitrageur advertises signal $\hat{\theta}_i = 1$. If advertising is successful (S), 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 (S) and at $t = 2$ the actual return is high $(\theta_i = 1)$, then the investors' belief becomes

$$\mu_2(1, 1) = \frac{\mu_1(S)}{\mu_1(S) + (1 - \mu_1(S))\pi} > \mu_1(S). \tag{20}$$

Since the high-skill arbitrageur always rightly identifies an undervalued asset $(\hat{\theta}_i = \theta_i)$, the numerator of (20) 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 $\mu_2(1, 1) > \mu_1(S)$ indicates that when investors observe that the arbitrageur's advertisement was correct, they revise their belief upward. Somewhat abusing the notation, for brevity we denote $\mu_2(1, 1) = \mu_2(1)$, $\mu_2(1, 0) = \mu_2(0)$ and $\mu_2(\emptyset, 1) = \mu_2(\emptyset, 1) = \mu_2(\emptyset)$.

In order to state the arbitrageur's problem at $t = 0$, it remains to specify the probability distribution of the possible realizations of investors' beliefs from the arbitrageur's point of view. This distribution depends on his type. The high-skill arbitrageur observes $\hat{\theta}_i = \theta_i$ and exerts effort e_H, hence $\Pr[\mu_2(0)] = 0$, $\Pr[\mu_2(\emptyset)] = 1 - e_H$ and $\Pr[\mu_2(1)] = e_H$ (recalling that by Assumption 8 $a = 1$). For the low-skill arbitrageur, $\hat{\theta}_i$ is independent of θ_i and the distribution is different: $\Pr[\mu_2(0)] = e_L(1-\pi)$, $\Pr[\mu_2(\emptyset)] = 1 - e_L$ and $\Pr[\mu_2(1)] = e_L\pi$.

Using (17) we express the price reaction to advertising in period $t = 1$ as $\Delta_1 = \mu_1(S)(1-\pi)$. Now we can state the arbitrageur's maximization problem at $t = 0$ taking into account the effect of investors' beliefs on his future payoff V' after $t = 3$:

$$\max_{e_\tau \in [0,1]} \frac{e_\tau \Delta_1}{\pi_i} w + w - e_\tau^2/2 + E[V'|e_\tau, \tau],\ \tau \in \{L, H\}. \tag{21}$$

For the high-skill arbitrageur we have $E[V'|e_H, H] = e_H V'(\mu_2(1)) + (1-e_H)V'(\mu_2(\emptyset))$, and for the low-skill $E[V'|e_L, L] = e_L[\pi V'(\mu_2(1)) + (1-\pi)V'(\mu_2(0))] + (1-e_L)V'(\mu_2(\emptyset))$. Problem (21) is convex and has a solution. With pessimistic beliefs, there exist equilibria without advertising at $t = 0$. We concentrate on equilibria in which at least one type of arbitrageurs does advertise at $t = 0$.

Proposition 6. *An equilibrium with advertising exists. In this equilibrium the high-skill arbitrageur advertises more at $t = 0$ than the low-skill one: $e_H^* > e_L^*$.*

The proof is in the appendix. Intuitively, the high-skill arbitrageur is confident of his information and knows that successful advertising will improve both his reputation and his future expected payoff at $t = 2$. Therefore, he exerts high effort. The low-skill arbitrageur is not confident of his information and if he advertises successfully will be proven wrong at $t = 2$ with some probability. If he turns out to be wrong, his reputation collapses and his expected payoff drops. Anticipating this, he advertises less than the high-skill arbitrageur.

As mentioned, the interaction from $t = 3$ onward is equivalent to the static model, simply replacing investors' belief about the arbitrageur's type with μ_2. Lemma 4 describes how the arbitrageur's reputation evolves in equilibrium:

Lemma 4. *In equilibrium the arbitrageur's reputation improves between $t = 0$ and $t = 1$ if advertising succeeds, and deteriorates if it fails: $\mu_1(F) < \mu < \mu_1(S)$. If advertising succeeds, reputation improves further at $t = 2$ if the asset's actual return is high and drops to zero if the return is low: $\mu_2(1) > \mu_1(S)$ and $\mu_2(0) = 0$.*

The proof is in the appendix. The result is intuitive. In equilibrium, the high-skill arbitrageur exerts greater advertising effort, so successful advertising is a noisy signal of the arbitrageur's type: investors revise their beliefs about his type upward. If advertising fails, so that investors do not learn anything from the arbitrageur, they revise their beliefs downward and the arbitrageur's reputation drops. In this case, actual returns at $t = 2$ are not informative about the arbitrageur's type because investors cannot compare them with his announcement. When successful advertising ($\hat{s}_i = \hat{\theta}_i = 1$) at $t = 1$ is confirmed by the actual return $\theta_i = 1$ at $t = 2$, investors revise their beliefs up. When instead successful advertising ($\hat{s}_i = \hat{\theta}_i = 1$) at $t = 1$ is shown to be wrong at $t = 2$ by the actual asset return ($\theta_i = 0$), investors realize that the arbitrageur is low-skill and his reputation drops to zero.

The advertisement of the high-skill arbitrageur is more likely to coincide with the actual return than that of the low-skill arbitrageur, so that:

Remark 2. *The reputation of the high-skill arbitrageur is more likely than that of the low-skill one to increase from $t = 1$ to $t = 2$.*

We now turn to the price reaction to advertising. Equation (17) determines the equilibrium price reaction to successful advertising for a given level of reputation. At $t = 1$ if advertising succeeds $\Delta_1 = \mu_1(S)(1-\pi)$, which together with Lemma 4 ($\mu_1(S) > \mu$) proves:

Proposition 7. *With reputation-building, the equilibrium price reaction at $t = 1$ is greater than in the static model: $\Delta_1^* > \Delta^*$.*

The result is intuitive. Reputation-building means that in equilibrium the high-skill arbitrageur exerts more effort than the low-skill one (Proposition 6). Successful advertising boosts the arbitrageur's reputation (Lemma 4) and, thus induces investors to react more strongly to successful advertising than in the static model, where both types of arbitrageur exert the same effort and successful advertising has no reputational effect.

It is interesting to study how the price reaction to advertising evolves over time as a function of arbitrageurs' reputation. Consider how prices at $t = 4$ react to advertising at $t = 3$. If advertising succeeded at $t = 1$ but the actual return was low ($\theta_i = 0$) at $t = 2$, then $\mu_2(0) = 0$ and the price does not react to new advertising at $t = 3$: $\Delta_4(0) = \mu_2(0)(1-\pi) = 0$. If advertising succeeded at $t = 1$ and the actual return was high ($\theta_i = 1$) at $t = 2$, then by Lemma 4 the investors' equilibrium belief at $t = 2$ rises to $\mu_2(1)$ and the price reaction is $\Delta_4(1) = \mu_2(1)(1-\pi)$. If advertising failed at $t = 1$, then at $t = 4$ the investors' equilibrium belief is $\mu_2(\emptyset)$ and the price reaction is $\Delta_4(\emptyset) = \mu_2(\emptyset)(1-\pi)$. Comparing these price reactions with the price reaction at $t = 1$, δ_1^*, yields:

Proposition 8. *The price impact of advertising increases over time if it proves to be accurate and diminishes if it is wrong. More precisely, 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 (and smaller than in the static model) if advertising fails, and least if successful advertising at $t = 1$ is belied by the asset's return at $t = 2$: $\Delta_4^*(1) > \Delta^* > \Delta_4(\emptyset) > \Delta_4(0) = 0$.*

The proof follows directly from Proposition 6 and Lemma 4. Naturally, 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 Lemma 1. If $e_j^* = 0$, then $q_j = \min[a_j e_j, 1] = 0$. Consider $e_i > 0$ for some $i \in M$ $q_i < 1$. Fix $\mathbf{y}^*$ and e_k^*, $k \neq i$. To see that $q_i < 1$ suppose instead that $q_i = 1$ and $e_i = \frac{1}{a_i}$: then the first order condition with respect to e_i would require

$$a_i e_i E[V_e(\sum_k \tilde{r}_k y_k, 1/a_i + \sum_{k \neq i} e_k)] \geq -a_i E[V(\frac{y_i}{\pi_i} + \sum_{k \neq i} \tilde{r}_k y_k, 1/a_i + \sum_{k \neq i} e_k)].$$

First, $\tilde{r}_i \leq \frac{1}{\pi_i}$, $\pi_i \geq \underline{\pi}$ and $\sum_k y_k \leq w$ implies $\sum_k \tilde{r}_k y_k \leq \frac{w}{\underline{\pi}}$. Second, $a_i \leq 1$ implies $\frac{1}{a_i} + \sum_{k \neq i} e_k \geq 1$. Together with $V_{ce} \geq 0$ and $V_{ee} \leq 0$ this implies that the left-hand side is less than $V_e(\frac{w}{\underline{\pi}}, 1)$. Together with $V_c > 0$ and $V_e < 0$ this implies that the right-hand side is greater than $-V(\frac{w}{\underline{\pi}}, 1)$, which contradicts Assumption 5. Thus $q_i < 1$, $i \in M$. QED.

Proof of Proposition 1. Consider a solution $\mathbf{y}^*$, $\mathbf{e}^*$ to (4). Since $V_e(c, 0) = 0$, $\gamma_i < 1$ and $a_i > 0$ for any $i \in M$ we must have $e_i^* > 0$ for some $i \in M$. First, 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 $e_j = \hat{e} - e_i$.

Lemma 1 implies that $q_i < 1$, $q_j < 1$. 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 $e_j = \hat{e} - e_i$. Substitute for e_j in (3). Suppose $e_j^* > 0$. The first order condition for an interior solution requires $\frac{\partial E(V|\bar{y}, \bar{e})}{\partial e_i}|_{e_j = \hat{e} - e_i} = 0$.

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

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

We will show that if $V(c, e)$ is concave in c, then $\frac{\partial^2 E(V|\bar{y}, \bar{e})}{\partial^2 e_i}|_{e_j = \hat{e} - e_i} > 0$, which means that at the optimum either e_i^*, or e_j^* should be zero. First, note that $\frac{\partial^2 E(V|\bar{y}, \bar{e})}{\partial^2 e_i}|_{e_j = \hat{e} - e_i} \geq 0$ is

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	$\frac{1}{2}(1-t_i)(1-t_j)$

equivalent to

$$\begin{aligned}\frac{1}{2}E[V(y_i r_i^H + y_j \rho_j + \sum_{k\neq i,j} \tilde{r}_k y_k)] + \frac{1}{2}E[V(y_i \rho_i + y_j r_j^H + \sum_{k\neq i,j} \tilde{r}_k y_k)] \geq \\ \frac{1}{2}E[V(y_i r_i^H + y_j r_j^H + \sum_{k\neq i,j} \tilde{r}_k y_k)] + \frac{1}{2}E[V(y_i \rho_i + y_j \rho_j + \sum_{k\neq i,j} \tilde{r}_k y_k)].\end{aligned} \tag{22}$$

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 $\tilde{z} = \sum_{k\neq i,j} \tilde{r}_k y_k$ which is independent of the returns on assets i, j. Note that the right-hand side of (22) 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 (22) 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 $\tilde{x}_L$.

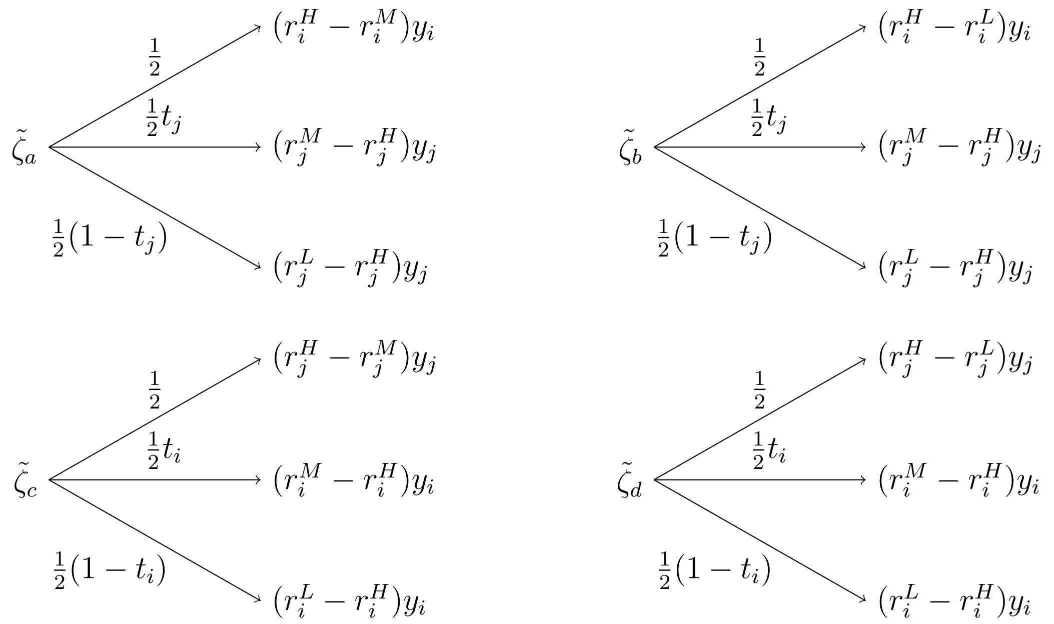
First, note that returns on assets $k \neq i, j$ do not matter for the comparison. Now, consider assets i and j. The table below describes possible realizations of monetary returns from assets i and j in lotteries $\tilde{x}_L$ and $\tilde{x}_R$ with corresponding probabilities.

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:

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

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 4: 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. $V(c, e)$ is concave in c, then $\frac{\partial^2 E(V|\bar{y},\bar{e})}{\partial^2 e_i}|_{e_j=\hat{e}-e_i} > 0$. 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. $V(c, e)$ is linear in 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 $q_i < 1$, $q_j < 1$ from Lemma 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. QED.

Proof of Lemma 2. When advertising is not possible, the arbitrageur's portfolio choice $\mathbf{y}$ must satisfy his resource constraint $\sum_i y_i = w$ and maximize

$$
\begin{aligned}
&t^2 E[V(y_k r^M + y_i r^M + \sum_{j\neq i,k} y_j\rho_j)] + t(1-t)E[V(y_k r^M + y_i r^L + \sum_{j\neq i,k} y_j\rho_j)]+ \\
&(1-t)tE[V(y_k r^L + y_i r^M + \sum_{j\neq i,k} y_j\rho_j)] + (1-t)^2 E[V(y_k r^L + y_i r^L + \sum_{j\neq i,k} y_j\rho_j)]
\end{aligned}
\tag{23}
$$

Table 2: Investments y_i' and y_j.

return on investment	probability for $y_i' = y$, $e_i' = e$	probability for $y_j = y$, $e_j = e$.
$r^H y$	$a_i e$	$a_j e$
$r^M y$	$(1 - a_i e)t$	$(1 - a_j e)t$
$r^L y$	$(1 - a_i e)(1 - t)$	$(1 - a_j e)(1 - t)$

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 (23) 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

$$\begin{aligned}
&t^2 E[V'(y_k r^M + (\bar{y} - y_k) r^M + \sum_{j \neq i,k} y_j^* \rho_j)](r^M - r^M) + \\
&t(1-t) E[V'(y_k r^M + (\bar{y} - y_k) r^L + \sum_{j \neq i,k} y_j^* \rho_j)](r^M - r^L) + \\
&(1-t)t E[V'(y_k r^L + (\bar{y} - y_k) r^M + \sum_{j \neq i,k} y_j^* \rho_j)](r^L - r^M) + \\
&(1-t)^2 E[V'(y_k r^L + (\bar{y} - y_k) r^L + \sum_{j \neq i,k} y_j^* \rho_j)](r^L - r^L) = 0.
\end{aligned} \tag{24}$$

As the first term and the last term of the left-hand side of (24) are zero, equation (24) becomes: $E[V'(y_k r^M + (\bar{y} - y_k) r^L + \sum\limits_{j \neq i,k} y_j^* \rho_j)] = E[V'(y_k r^L + (\bar{y} - y_k) r^M + \sum\limits_{j \neq i,k} y_j^* \rho_j)]$, 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. Recall that, by assumption, when advertising effort is zero, its marginal cost is zero: $V_e(c, 0) = 0$. First, the arbitrageur must advertise the asset with the greatest advertisability $i = \arg\max\limits_{k \in M} a_k$. Suppose otherwise $e_i = 0$ and $e_j = e > 0$ 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 $e_j' = e_i$. 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:

Since $i = \arg\max_{k \in M} a_k$, it must be that $e_i > 0$.

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 2.

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

$$\begin{aligned}
E[V|\mathbf{y}, e] = & ea_i t E[V(y_i r^H + y_j r^M + \sum_{k \neq i,j} y_k \rho_k, e)] + \\
& ea_i (1-t) E[V(y_i r^H + y_j r^L + \sum_{k \neq i,j} y_k \rho_k), e] + \\
& (1 - ea_i) t^2 E[V(y_i r^M + y_j r^M + \sum_{k \neq i,j} y_k \rho_k, e)] + \\
& (1 - ea_i) t(1-t) E[V(y_i r^M + y_j r^L + \sum_{k \neq i,j} y_k \rho_k, e)] + \\
& (1 - ea_i)(1-t) t E[V(y_i r^L + y_j r^M + \sum_{k \neq i,j} y_k \rho_k, e)] + \\
& (1 - ea_i)(1-t)^2 E[V(y_i r^L + y_j r^L + \sum_{k \neq i,j} y_k \rho_k, e)].
\end{aligned} \tag{25}$$

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 (25) over y_i given the constraint $y_j = \bar{y} - y_i$. The first order necessary condition with respect to y_i is:

$$\begin{aligned}
& ea_i t E[V'(y_i r^H + y_j r^M + \sum_{k \neq i,j} y_k \rho_k, e)](r^H - r^M) + \\
& ea_i (1-t) E[V'(y_i r^H + y_j r^L + \sum_{k \neq i,j} y_k \rho_k), e](r^H - r^L) + \\
& (1 - ea_i) t^2 E[V'(y_i r^M + y_j r^M + \sum_{k \neq i,j} y_k \rho_k, e)](r^M - r^M) + \\
& (1 - ea_i) t(1-t) E[V'(y_i r^M + y_j r^L + \sum_{k \neq i,j} y_k \rho_k, e)](r^M - r^L) + \\
& (1 - ea_i)(1-t) t E[V'(y_i r^L + y_j r^M + \sum_{k \neq i,j} y_k \rho_k, e)](r^L - r^M) + \\
& (1 - ea_i)(1-t)^2 E[V'(y_i r^L + y_j r^L + \sum_{k \neq i,j} y_k \rho_k, e)](r^L - r^L) = 0.
\end{aligned} \tag{26}$$

This reduces to

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

The LHS is positive for any $e > 0$. 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 Corollary 1. Assume that $Lw < \underline{\pi}$ holds. Given that $a_j > 0$, $\gamma_j < 1$ and $\pi_j < 1$, function $V_j(L)$ increases with L and $V_j(L) \to \infty$ if $L \to \infty$, while V^a_{-j} is bounded from above. It follows that for any $j \in M$ one can find $\underline{L}(j) < \infty$ such that for $L \geq \underline{L}(j)$ one has $V_j(L) \geq V^a_{-j}$. Expressions for $V_j(L)$ and V^a_{-j} are derived assuming that $w < \underline{\pi}/L$. Hence, if the latter condition is verified and if $L \geq \underline{L}(j)$ an equilibrium with all arbitrageurs advertising asset j exists. QED.

Proof of Lemma 4. According to Proposition 6 $e^*_H > e^*_L$, which together with (13) and (15) implies $\mu_1(F) < \mu < \mu_1(S)$. If advertising fails, investors learn nothing about the arbitrageur when they see actual returns at $t = 2$: $\mu_2(\emptyset) = \mu_1(F)$. If advertising succeeds but is contradicted consistent by the actual returns, the arbitrageur is necessarily low-skill: $\mu_2(0) = 0$. Finally, if advertising succeeds and is consistent with the actual returns, equation (20) implies $\mu_2(1) > \mu_1(S)$. QED.

Proof of Proposition 6. First, let us prove that an equilibrium with $e^*_H > 0$ exists. Consider a low-skill arbitrageur who maximizes the objective function (21). In case of an interior solution, his optimal effort is given by $e_L = f(e_L, e_H)$ where

$$f(e_L, e_H) = \mu_1(S)(1-\pi) + \pi V'(\mu_2(1)) + (1-\pi)V'(\mu_2(0)) - V'(\mu_2(\emptyset)).$$

Substituting for V', $\mu_2(1)$, $\mu_2(0)$, $\mu_2(\emptyset)$ and using Assumption 8 this expression becomes

$$f(e_L, e_H) = (1-\pi)\frac{\mu e_H}{\mu e_H + (1-\mu)e_L} +$$
$$\frac{(1-\pi)^2}{2}\left[\pi\left(\frac{\mu e_H}{\mu e_H + (1-\mu)\pi e_L}\right)^2 - \left(\frac{\mu(1-e_H)}{\mu(1-e_H) + (1-\mu)(1-e_L)}\right)^2\right].$$

Hence, the optimal effort of the low-skill arbitrageur is

$$e_L^* = \begin{cases} 0 & \text{if } f(0, e_H) < 0, \\ 1 & \text{if } f(1, e_H) > 1, \\ f(e_L^*, e_H) & \text{otherwise.} \end{cases}$$

Take an arbitrary small $\epsilon > 0$. Function $f(e_L, e_H)$ is continuous for $e_H \in [\epsilon, 1]$, decreases with e_L and increases with e_H, so that the first order condition defines a continuous increasing function $e_L^* = \underline{e}_L(e_H^*)$ for $e_H^* \in [\epsilon, 1]$.

Now consider a high-skill arbitrageur who maximizes the objective function (21). In case of an interior solution, his optimal effort is given by $e_H = g(e_L, e_H)$ where

$$g(e_L, e_H) = \mu_1(S)(1 - \pi) + V'(\mu_2(1)) - V'(\mu_2(\emptyset)).$$

Upon substituting, this expression becomes

$$g(e_L, e_H) = (1 - \pi)\frac{\mu e_H}{\mu e_H + (1 - \mu)e_L} + \frac{(1 - \pi)^2}{2}\left[\left(\frac{\mu e_H}{\mu e_H + (1 - \mu)\pi e_L}\right)^2 - \left(\frac{\mu(1 - e_H)}{\mu(1 - e_H) + (1 - \mu)(1 - e_L)}\right)^2\right].$$

Hence, the optimal effort of the high-skill arbitrageur is

$$e_H^* = \begin{cases} 0 & \text{if } g(e_L, 0) < 0, \\ 1 & \text{if } g(e_L, 1) > 1, \\ g(e_L, e_H^*) & \text{otherwise.} \end{cases}$$

Given that $g(e_L, e_H)$ is continuous for $e_H \in [\epsilon, 1]$ and decreases with e_L, the first order condition defines a continuous function $e_L^* = \bar{e}_L(e_H^*)$ for $e_H^* \in [\epsilon, 1]$.

Consider $e_H = 1$. Note that $\underline{e}_L(e_H) \leq \underline{e}_L(1) < 1$ because

$$\underline{e}_L(1) = f(\underline{e}_L(1), 1) \leq f(0, 1) = (1 - \pi) + \frac{(1 - \pi)^2}{2}\pi = (1 - \pi/2)(1 - \pi^2) < 1.$$

Two cases are possible. If $g(\underline{e}_L(1), 1) \geq 1$, then $e_H^* = 1$ and $e_L^* = \underline{e}_L(1)$ satisfy the equilibrium conditions, and an equilibrium with positive advertising effort exists.

If instead $g(\underline{e}_L(1), 1) < 1$, then $\underline{e}_L(1) > \bar{e}_L(1)$, because $g(e_L, e_H)$ decreases with e_L and $g(\bar{e}_L(1), 1) = 1$. The functions $\bar{e}_L(e_H^*)$ and $\underline{e}_L(e_H^*)$ are continuous for $e_H \in [\epsilon, 1]$, hence if $\underline{e}_L(\epsilon) < \bar{e}_L(\epsilon)$ for $\epsilon \to 0$, the functions intersect for some $e_H \in [\epsilon, 1]$ and an equilibrium with positive efforts exists. Consider $e_H = \epsilon \to 0$ and note that $\underline{e}_L(\epsilon) > \epsilon$. Indeed,

$f(\epsilon,\epsilon) = \mu(1-\pi) + \frac{(1-\pi)^2}{2}[\pi\left(\frac{\mu}{\mu+(1-\mu)\pi)}\right)^2 - \mu^2] > \mu(1-\pi)[1-\frac{(1-\pi)^2}{2}\mu] > \mu(1-\pi)/2 > \epsilon$ for small enough ϵ. Since f decreases with e_L, the solution to $\underline{e}_L(\epsilon) = f(\underline{e}_L(\epsilon), \epsilon)$ must have $\underline{e}_L(\epsilon) > \epsilon$. Note that $\overline{e}_L(\epsilon)$ solves $\epsilon = g(\overline{e}_L(\epsilon), \epsilon)$. Let us now prove $\underline{e}_L(\epsilon) < \overline{e}_L(\epsilon)$, which is equivalent to $g(\underline{e}_L(\epsilon), \epsilon) > \epsilon$ because $g(e_L, e_H)$ decreases with e_L. This is true because $\underline{e}_L(\epsilon) = f(\underline{e}_L(\epsilon), \epsilon)$ and

$$g(\underline{e}_L(\epsilon), \epsilon) - f(\underline{e}_L(\epsilon), \epsilon) = \frac{(1-\pi)^3}{2}\left(\frac{\mu e_H}{\mu e_H + (1-\mu)\pi e_L}\right)^2 > 0 > \epsilon - \underline{e}_L(\epsilon).$$

Therefore, $\underline{e}_L(\epsilon) < \overline{e}_L(\epsilon)$. Given that we consider the case $\underline{e}_L(1) > \overline{e}_L(1)$ an equilibrium with positive efforts exists with $e_H^* \in [\epsilon, 1]$.

Finally, we must prove that $e_H^* > e_L^*$. In an equilibrium with advertising $e_L^* = f(e_L^*, e_H^*) \in (0,1)$ as we have shown above. The high-skill arbitrageur chooses $e_H^* = \min[1, g(e_L^*, e_H^*)] > e_L^*$ because $g(e_L, e_H) > f(e_L, e_H)$ and $f(e_L^*, e_H^*) < 1$ in equilibrium. 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*.