THE JOURNAL OF FINANCE • VOL. LXIV, NO. 4 • AUGUST 2009

Presidential Address: Sophisticated Investors and Market Efficiency

JEREMY C. STEIN*

ABSTRACT

Stock-market trading is increasingly dominated by sophisticated professionals, as opposed to individual investors. Will this trend ultimately lead to greater market efficiency? I consider two complicating factors. The first is crowding—the fact that, for a wide range of "unanchored" strategies, an arbitrageur cannot know how many of his peers are simultaneously entering the same trade. The second is leverage—when an arbitrageur chooses a privately optimal leverage ratio, he may create a fire-sale externality that raises the likelihood of a severe crash. In some cases, capital regulation may be helpful in dealing with the latter problem.

IN THE LAST 20 YEARS or so, there have been profound changes in the way that money is managed. One indicator of these changes is the rapid growth of the hedge fund industry, whose assets on a global basis have gone from \$39 billion at year-end 1990 to \$1.93 trillion as of the second quarter of 2008.[1] Hedge funds are commonly thought of as the prototypical sophisticated investors, for a couple of reasons. First, many of their investment strategies are based on extensive quantitative modeling, much of which has its roots in academic research in finance.[2] Second, hedge funds often implement these strategies in an aggressively leveraged fashion.

The growth of hedge funds is part of a broader trend toward professional asset management. French (2008) documents that, in the stock market, individual investors have been largely supplanted by institutions. Direct individual ownership of U.S. equities, which was 47.9% in 1980, fell to 21.5% by 2007. At

*Jeremy C. Stein is from Harvard University and NBER. Presidential Address delivered to the American Finance Association, San Francisco, January 4, 2009. I am indebted to Sam Hanson for outstanding research assistance. I also thank John Campbell, Harrison Hong, Robin Greenwood, Borja Larrain, Andrei Shleifer, Erik Stafford, Dimitri Vayanos, Ivo Welch, and seminar participants at Harvard for helpful comments and suggestions.

[1] These numbers are from Hedge Fund Research (HFR).

[2] Of particular interest to many "quant" funds has been the large body of empirical work documenting various patterns of predictability in stock returns, many of which are strikingly robust over time and across countries. A partial list includes: the value-glamour effect (Fama and French (1992), Lakonishok, Shleifer, and Vishny (1994)); medium-run momentum (Jegadeesh and Titman (1993)); post-earnings announcement drift (Bernard and Thomas (1989, 1990)); and the low returns to firms with high levels of past accruals, equity issuance, or asset growth (Sloan (1996), Daniel and Titman (2006), and Cooper, Gulen, and Schill (2008), respectively). See Asness, Moskowitz, and Pedersen (2008) for a recent synthesis.

the same time, turnover has exploded, reaching 215% in 2007.[3] If one adopts the view that individuals are naïve investors while institutions are rational arbitrageurs, these data would seem to suggest that we are converging to a world in which the smart-money players trade intensively with one another, with the dumb money playing a much-diminished role.

My goal in this paper is to explore the consequences of these developments for market efficiency. I pose the following conceptual question. Imagine a market in which there are both naïve investors with biased expectations and fully rational arbitrageurs. Now let the capital controlled by the latter group grow increasingly large relative to that of the former. Is it the case that the market is necessarily made more efficient, in the sense that prices on average wind up closer to fundamental values, and nonfundamental sources of volatility become less important?

At first glance, it might appear that the answer to this question would have to be yes. Basic economic logic suggests that as more money is brought to bear against a given trading opportunity, any predictable excess returns must be reduced and eventually eliminated. This is nothing more than a zero-profit condition that must be met when the risk tolerance of the arbitrageurs goes to infinity. A corollary would seem to be that, to the extent that predictable excess returns reflect an underlying inefficiency, their elimination must go hand in hand with prices being pushed closer to fundamentals.

While the zero-profit intuition is certainly on target—enough money chasing a given pattern in returns will necessarily eliminate that pattern—I argue that the corollary does not follow. In particular, I argue that it need not be the case that the elimination of predictability is associated with either an on-average narrowing of the gap between prices and fundamentals, or a reduction in nonfundamental volatility. Thus, while a larger number of sophisticated arbitrageurs certainly makes life more competitive and less profitable for the arbitrageurs themselves, it need not make the world a better place for those who look to asset prices to provide a reliable reflection of underlying fundamental values.

Where, then, does the simple intuition about competition and market efficiency go wrong? In the most general terms, complications arise when, in the process of pursuing a given trading strategy, arbitrageurs inflict negative externalities on one another. In this paper, I model two distinct mechanisms by which such externalities are created. The first has to do with what might be termed a "crowded-trade" effect. For a broad class of quantitative trading strategies, an important consideration for each individual arbitrageur is that he cannot know in real time exactly how many others are using the same model and taking the same position as him. This inability of traders to condition their behavior on current market-wide arbitrage capacity creates a coordination problem and, as I show further, in some cases can result in prices being pushed further away from fundamentals.

A second way in which arbitrageurs inflict externalities on one another is through their leverage decisions. If two traders follow the same set of signals

[3] The numbers in this paragraph are taken from French (2008).

and buy the same stocks using leverage, then if one is hit with a negative shock—say, losses in an unrelated part of his portfolio—he will be forced to liquidate some of the commonly held stocks to meet margin calls, potentially creating a fire-sale effect in prices and inflicting losses on the other trader, thereby generating another round of liquidations and price declines.

The crowded-trade issue has not, to my knowledge, received much formal research attention. By contrast, the problems associated with leveraged arbitrageurs have been extensively studied, by, among others, Shleifer and Vishny (1992,1997), Kyle and Xiong (2001), Gromb and Vayanos (2002), Morris and Shin (2004), Allen and Gale (2005), and Brunnermeier and Pedersen (2008). Indeed, the fire-sale mechanism analyzed in these papers occupies a central place in accounts of the demise of Long-Term Capital Management (LTCM) in 1998, and more recently, the "quant crisis" of August 2007.

Where this paper differs from much of the previous literature is in analyzing the consequences of both crowding and leverage in a setting in which arbitrageurs (i) have rational expectations; (ii) make optimizing leverage decisions ex ante; and (iii) have access to a potentially infinite amount of equity capital. Thus, with respect to the crowded-trade issue, I assume that each arbitrageur makes an unbiased estimate of the number of others that are active in the market at any point in time. And with respect to leverage, arbs can choose not to borrow at all if they prefer not to face the risk of having to liquidate during a fire sale; moreover, they have access to enough capital in the aggregate that they can take positions of any size without having to resort to borrowing.

The usefulness of this approach can be illustrated by reference to the quant crisis. During the week of August 6, 2007, many popular quantitative strategies simultaneously experienced enormous negative returns—in several cases, the daily movements were on the order of 10 or more standard deviations relative to historical norms. The emerging consensus about this episode is that the proximate causes of the crisis included overcrowding and overleverage: Too many quant managers were invested in the same strategies, with too much leverage (see, for example, Khandani and Lo (2007)). One way to interpret this story is that it is about a set of one-time mistakes made by the quant managers: Given the rapid growth in this sector, quant managers grossly underestimated the total amount of money invested in their favorite strategies and then compounded this error by leveraging their positions to a degree that, at least in hindsight, seems excessive. This one-time mistake interpretation of the crisis yields an optimistic view of the future. It suggests that, having learned the errors of their ways, rational arbitrageurs will pull back from the most overcrowded strategies and reduce their leverage. And once they do so, they will go back to being a force for stability and market efficiency.

I do not dispute that there can be substantial learning in the wake of extreme events such as those of August 2007. However, my analysis implies that such learning is not necessarily a panacea. The problems associated with crowding and leverage that I identify persist even when arbitrageurs fully understand the structure of the world that they are operating in, arrive at unbiased estimates of all of the relevant parameters, and optimize their capital structures accordingly. Thus, any remaining inefficiencies cannot be said to be the result of one-time mistakes, but rather must be thought of as a more permanent part of the landscape.

This last observation suggests a potential role for regulatory policy. In particular, my analysis of arbitrageur capital structure makes it clear that privately chosen leverage ratios may be greater than socially optimal ones, implying that some form of capital regulation might improve market efficiency. At the same time, the analysis also points to some of the costs of such regulation, and highlights the difficulties associated with striking a proper balance.

In what follows, I develop two simple models to address the issues associated with crowding and arbitrageur leverage choice. These models are presented in Section I and II, respectively. Section III concludes the paper.

I. The Crowded-Trade Problem

A. Overview

I begin by exploring a simple model of the crowded-trade problem. This model is built on two key premises. First, while there is potentially a large amount of arbitrage capacity available to take on mispricings of any given stock at any point in time, *no individual arbitrageur knows exactly how much is available*. This uncertainty could reflect each arbitrageur's imperfect information as to: (i) the number of other players who might be pursuing a particular trading strategy; (ii) their current capital and liquidity positions; or (iii) the nature of their alternative investment opportunities.

Second, the trading strategy in question is one with *no fundamental anchor*: Arbitrageurs do not base their demand on an independent estimate of fundamental value. As a result, *their demand for an asset may be a nondecreasing function of the asset's price*.[4] Strategies of this type are common in practice, and include many in which demand is independent of price, for example: (i) buying the stocks of firms with low values of accruals, equity issues, or asset growth; (ii) buying small-cap stocks in December; or (iii) buying stocks that are expected to be added to a widely tracked index. There are also strategies in which demand is an increasing function of price, such as those used to exploit momentum in stock returns. In the model below, I focus on a post-earnings announcement drift (PEAD) strategy that has a momentum-like flavor: Arbs buy when returns on the earnings announcement day are positive, and sell when they are negative.

Trading strategies like these can certainly be profitable. However, they make the market vulnerable to the effects of overcrowding. Consider a strategy of buying low-accrual stocks without regard to price. This strategy may earn positive returns for an arbitrageur on average, to the extent that it exploits the tendency of less rational investors to undervalue such stocks. However, if an unexpectedly large number of other arbs suddenly adopt the same strategy, there is no

[4] In this sense, the arbitrageurs in my model are similar to the rational but uninformed investors in Grossman and Stiglitz (1980).

price-based mechanism to mediate the congestion that arises, and these stocks may become overvalued.

This feature of the model distinguishes it from many standard treatments of arbitrage, including DeLong et al. (1990a), where arbitrageurs observe fundamental value perfectly and thus have demands that are a simple function of the gap between fundamentals and price. The DeLong et al. setting corresponds most closely to spread trades, such as those involving "Siamese twin" stocks (Froot and Dabora (1999)), where the spread itself is easily observed. In a spread-trade situation, if an unexpectedly large number of arbs show up, the spread narrows and each individual arb can adjust his demand accordingly. In other words, the price mechanism mediates congestion, and there is no danger of the market becoming overcrowded with arbitrageurs.

B. The Trading Opportunity: Some Investors Underreact to News

There is a stock that pays a terminal dividend at time 2 of $V = F + \varepsilon$. The disturbance term ε is normally distributed with a mean of zero and a variance of 1, and it is realized at time 2. The fundamental F is also normally distributed with a variance of 1. As of time 0, the ex ante expectation of F is 0; the supply of the stock is also equal to 0. So the time-0 price is given by $P_0 = 0$.

At time 1 there is a news release, which can be thought of as an earnings announcement. This news release is seen by a group of investors that, following Hong and Stein (1999), I label the "newswatchers." However, while the newswatchers are good at paying attention to fundamental information, they do not process it in an unbiased fashion—rather, they underreact to it. That is, the newswatchers observe the realization of F, but use this data to form a biased expectation of V, $E^n(V) = F(1 - \delta)$, where δ, which lies between 0 and 1, measures the extent of their underreaction; higher values of δ are associated with a more extreme bias on the part of newswatchers.

Newswatchers have mean-variance preferences and unit aggregate risk tolerance. It follows that if they are the only players in the market, the time-1 price P_1, as well as the return from time 0 to time 1, R_1 (defined simply as the change in price), are both given by

$$P_1 = R_1 = F(1 - \delta). \tag{1}$$

The return from time 1 to 2 is given by

$$R_2 = V - P_1 = \delta F + \varepsilon = \frac{\delta R_1}{(1 - \delta)} + \varepsilon, \tag{2}$$

and the expected return over the same interval is given by

$$E(R_2 \mid R_1) = \frac{\delta R_1}{(1 - \delta)}. \tag{3}$$

Thus, there is a drift-like pattern in returns: A high return when news is released at time 1 forecasts further high returns from time 1 to 2, and more so when δ is large. This type of continuation in returns is due to the underreaction of the newswatchers to fundamental information. And clearly, it creates an arbitrage opportunity.

C. The Arbitrage Response

Next, I introduce a set of rational arbitrageurs who also have mean-variance preferences and who collectively have an aggregate risk tolerance of n. I then consider how the market outcome depends on: (i) whether or not the arbs observe the fundamental F and (ii) whether there is any uncertainty about total arbitrage capacity n.

C.1. Arbs Have Perfect Information

Suppose first that arbitrageurs observe the fundamental F, which they use to form an unbiased estimate of V. Newswatcher demand at time 1 is given by $D^n = F(1-\delta) - P_1$. Arbitrageur demand at time 1 is given by $D^a = n(F - P_1)$. Note that the arbs condition their demand directly on F; this is a case where their trading strategy *does* have a fundamental anchor, and where demand is therefore a decreasing function of price.

Setting total demand equal to the supply of 0, we have

$$P_1 = F - \frac{\delta F}{(1+n)}. \tag{4}$$

In this simple setting, more arbitrage always pushes prices closer to fundamentals: As arbitrage capacity n increases, the underreaction bias in prices is monotonically reduced.

C.2. Arbs Do Not Observe Fundamentals

I now consider what happens when arbitrageurs do not observe the news release F at time 1, that is, when they trade without being able to condition on a fundamental anchor. The interpretation is that while the nonarbitrage segment of the investment community (as represented by the newswatchers) makes some mistakes, as a group they have access to some valid information about fundamentals that the arbs do not.

Even though they cannot condition on F, the arbs still understand the structure of the model, and are aware of the bias that the newswatchers have in reacting to news. It follows that the arbs can make a profit by using a technical trading strategy, where their demand takes the form

$$D^a = n\phi P_1 = n\phi R_1. \tag{5}$$

Here, as in Hong and Stein (1999), the parameter ϕ is determined endogenously by optimization on the part of the arbitrageurs. As part of this optimization, arbs make an inference about future expected returns based on the current price P_1 (or equivalently, the most recent return R_1). This demand yields the following time-1 price

$$P_1 = F\frac{(1-\delta)}{(1-n\phi)}. \tag{6}$$

If both total arbitrage capacity n as well as newswatcher bias δ are fixed constants, it turns out that this case yields the same outcome as the one in which the arbs observe F directly. One way to understand this is to look at equation (6), and to note that if both n and δ are known by the arbs, they can infer F based on the price P_1. Given this inference, the arbs effectively have full information and hence we must be back to the full-information solution where $P_1 = F - \delta F/(1+n)$.

Although the outcome for prices is the same, it should be emphasized that it is implemented differently when arbitrageurs do not observe F. It can be easily shown that in equilibrium, $\phi = \delta/(1+n-\delta) > 0$. The arbs now play a PEAD-type strategy whereby they buy those stocks that have had positive announcement returns, that is, their individual demands are an increasing function of price, rather than a decreasing function as in the case where they condition directly on fundamentals. This distinction is of no consequence for equilibrium prices when aggregate arbitrage capacity is fixed, but is of crucial importance if we further introduce uncertainty on this dimension.

C.3. Uncertainty about Total Arbitrage Capacity

Suppose that we can write $n = N\theta$, where N is the expected arbitrage capacity at a given point in time, and θ is a random variable with a mean of one, independent of both F and ε, distributed on the interval $[\theta_L, \theta_H]$, with $\theta_L \geq 0$. This formulation implies that as the expected scale of the arbitrage sector (given by N) goes up, so does uncertainty about the exact amount of arbitrage capital that will be deployed on any given trade.

I continue to assume that, as before, each individual arb has a linear demand of the form $d^a = \phi R_1$. It should be pointed out that this linear demand function may no longer be unconditionally optimal when n is random—one might be able to improve on it with a nonlinear strategy. So I am effectively restricting arbs to *simple* strategies in what follows. I view this restriction as a reasonable approximation to real-world arbitrage behavior. Moreover, the parameter ϕ is still endogenous, so strategies are by definition optimal within the linear class, and arbs always make nonnegative expected profits.

Given that demands have the same form, equation (6) for the time-1 price continues to apply. What has changed, however, is that this price is no longer a summary statistic for the fundamental news F, since it now is also influenced by a second random variable, namely, n. Herein lies the danger for an individual arb who formulates his demand based on just the return at time 1. On the one hand, a large positive time-1 return could indicate a large realization of F, which assuming underreaction on the part of the newswatchers, would make the arb want to take a big long position at time 1. On the other hand, a large time-1 return could reflect a small positive realization of F, combined with an unexpectedly high level of arbitrage activity n. In this case, it is possible that the time-1 price has actually overshot its fundamental value, due to all the arbitrage buying. If so, any individual arb would be better off taking a short position.

To simplify the analysis, from this point forward I focus exclusively on the limiting risk-neutral case where expected arbitrage capacity N goes to infinity. To do so I define $\Phi = \lim_{N\to\infty}(\phi N)$. Here, Φ can be thought of as a measure of the expected trading intensity of the arbitrage sector as its risk tolerance becomes arbitrarily large. With this definition in hand, the time-1 price can be rewritten as

$$P_1 = F\frac{(1-\delta)}{(1-\theta\Phi)}. \tag{7}$$

It follows that the time-2 return is given by

$$R_2 = V - P_1 = \frac{F(\delta-\theta\Phi)}{(1-\theta\Phi)} + \varepsilon. \tag{8}$$

To solve the model, we need to solve for the equilibrium value of Φ, denoted by Φ^*. Observe that in the limiting risk-neutral case, expected returns to the linear trading strategy must be 0. This zero-profit condition implies that $\mathrm{cov}(R_1, R_2) = 0$.[5] Given the above expressions for returns, this means that, in equilibrium, we must have

$$E\left[\frac{(1-\delta)(\delta-\theta\Phi^*)}{(1-\theta\Phi^*)^2}\right] = 0. \tag{9}$$

If one specifies the distribution of the random variable θ, as well as a value of δ, equation (9) can be solved for the equilibrium trading intensity Φ^*. The solution to (9) is in general quite complicated, but it can be made relatively tractable if one picks a simple enough distribution for θ. Table I provides several illustrations, focusing on cases where θ has either a binomial or a uniform distribution. The analytical expressions for Φ^* in these cases are given in the Appendix. For each distribution, I compute Φ^*, $E(P_1/F)$, $\min(P_1/F)$, $\max(P_1/F)$, and $E\,|\,(P_1 - F)/F\,|$, experimenting with different values of: (i) the underreaction parameter δ and (ii) the variance of θ.

The key messages from Table I can be summarized as follows. First, the time-1 price underreacts to the fundamental F when there is a low realization of θ, that is, $\min(P_1/F) < 1$. Second, the price overreacts to F when there is a high realization of θ, that is, $\max(P_1/F) > 1$. These properties are intuitive, in that there needs to be a balance of under- and overreaction across realizations of θ in order for the zero-profit condition to be satisfied. If, instead, the price always

[5] To see this, note that demand is a linear function of R_1, which implies that the return to the strategy is proportional to R_1R_2. Setting expected returns equal to zero then yields $E(R_1R_2) = \mathrm{cov}(R_1R_2) = 0$.

Table I

Solutions to the Crowded-Trade Model for Various Parameter Values

The table displays: (i) equilibrium trading intensity; (ii) the mean of price to fundamentals; (iii) the minimum of price to fundamentals; (iv) the maximum of price to fundamentals; and (v) the mean absolute gap between price and fundamentals. These values depend on the underreaction parameter δ and the distribution of θ, which governs arbitrage capacity.

Panel A: Symmetric Binomial: $\theta = (1-h)$ with Prob $\frac{1}{2}$, $\theta = (1+h)$ with Prob $\frac{1}{2}$; $\mathrm{Var}[\theta] = h^2$

	$h = 0.5$	$h = 0.6$	$h = 0.7$	$h = 0.8$	$h = 0.9$
			Trading Intensity Φ^*		
$\delta = 0.1$	0.095	0.093	0.091	0.089	0.087
$\delta = 0.2$	0.180	0.174	0.168	0.161	0.155
$\delta = 0.3$	0.257	0.245	0.233	0.223	0.212
$\delta = 0.4$	0.326	0.308	0.291	0.276	0.263
$\delta = 0.5$	0.388	0.365	0.344	0.325	0.308
			$E(P_1/F)$		
$\delta = 0.1$	0.997	0.996	0.995	0.994	0.993
$\delta = 0.2$	0.988	0.984	0.981	0.977	0.973
$\delta = 0.3$	0.971	0.964	0.957	0.950	0.944
$\delta = 0.4$	0.945	0.933	0.923	0.915	0.907
$\delta = 0.5$	0.909	0.894	0.881	0.871	0.861
			$\mathrm{Min}(P_1/F)$		
$\delta = 0.1$	0.945	0.935	0.925	0.916	0.908
$\delta = 0.2$	0.879	0.860	0.842	0.827	0.813
$\delta = 0.3$	0.803	0.776	0.753	0.733	0.715
$\delta = 0.4$	0.717	0.684	0.657	0.635	0.616
$\delta = 0.5$	0.620	0.585	0.558	0.535	0.516
			$\mathrm{Max}(P_1/F)$		
$\delta = 0.1$	1.050	1.058	1.065	1.072	1.078
$\delta = 0.2$	1.097	1.109	1.119	1.127	1.134
$\delta = 0.3$	1.139	1.151	1.160	1.168	1.174
$\delta = 0.4$	1.173	1.183	1.189	1.194	1.197
$\delta = 0.5$	1.197	1.202	1.205	1.206	1.207
			$E\lvert(P_1–F)/F\rvert$		
$\delta = 0.1$	0.052	0.061	0.070	0.078	0.085
$\delta = 0.2$	0.109	0.124	0.138	0.150	0.161
$\delta = 0.3$	0.168	0.188	0.204	0.218	0.229
$\delta = 0.4$	0.228	0.249	0.266	0.279	0.291
$\delta = 0.5$	0.288	0.308	0.324	0.336	0.346

Panel B: Uniform: θ is Uniformly Distributed on $[(1-h), (1+h)]$; $\mathrm{Var}[\theta] = h^2/3$

	$h = 0.5$	$h = 0.6$	$h = 0.7$	$h = 0.8$	$h = 0.9$
			Trading Intensity Φ^*		
$\delta = 0.1$	0.098	0.097	0.097	0.096	0.095
$\delta = 0.2$	0.192	0.189	0.186	0.183	0.179
$\delta = 0.3$	0.281	0.275	0.268	0.260	0.253
$\delta = 0.4$	0.365	0.353	0.341	0.329	0.317
$\delta = 0.5$	0.440	0.423	0.405	0.389	0.373
			$E(P_1/F)$		
$\delta = 0.1$	0.999	0.999	0.998	0.998	0.997
$\delta = 0.2$	0.995	0.993	0.991	0.989	0.987

(*continued*)

Table I—*Continued*

Panel B: Uniform: θ is Uniformly Distributed on $[(1-h), (1+h)]$; $\mathrm{Var}[\theta] = h^2/3$					
	$h = 0.5$	$h = 0.6$	$h = 0.7$	$h = 0.8$	$h = 0.9$
$\delta = 0.3$	0.987	0.982	0.977	0.973	0.968
$\delta = 0.4$	0.971	0.962	0.953	0.945	0.936
$\delta = 0.5$	0.944	0.929	0.915	0.902	0.889
			$\mathrm{Min}(P_1/F)$		
$\delta = 0.1$	0.946	0.937	0.927	0.918	0.909
$\delta = 0.2$	0.885	0.866	0.847	0.830	0.815
$\delta = 0.3$	0.815	0.786	0.761	0.738	0.718
$\delta = 0.4$	0.734	0.699	0.668	0.642	0.620
$\delta = 0.5$	0.641	0.602	0.569	0.542	0.519
			$\mathrm{Max}(P_1/F)$		
$\delta = 0.1$	1.056	1.066	1.077	1.087	1.097
$\delta = 0.2$	1.124	1.148	1.170	1.191	1.212
$\delta = 0.3$	1.212	1.249	1.284	1.317	1.347
$\delta = 0.4$	1.324	1.378	1.427	1.471	1.511
$\delta = 0.5$	1.473	1.544	1.607	1.663	1.712
			$E\|(P_1 - F)/F\|$		
$\delta = 0.1$	0.027	0.032	0.037	0.042	0.047
$\delta = 0.2$	0.060	0.070	0.081	0.090	0.099
$\delta = 0.3$	0.099	0.115	0.130	0.144	0.156
$\delta = 0.4$	0.147	0.169	0.188	0.205	0.220
$\delta = 0.5$	0.206	0.233	0.256	0.275	0.292

Panel C: Skewed Binomial: $\theta = 1/p$ with Prob p, $\theta = 0$ with Prob $(1-p)$; $\mathrm{Var}[\theta] = (1-p)/p$					
	$p = 0.1$	$p = 0.2$	$p = 0.3$	$p = 0.4$	$p = 0.5$
			Trading Intensity Φ^*		
$\delta = 0.1$	0.041	0.059	0.071	0.079	0.085
$\delta = 0.2$	0.056	0.089	0.114	0.133	0.149
$\delta = 0.3$	0.064	0.109	0.146	0.176	0.203
$\delta = 0.4$	0.071	0.125	0.171	0.213	0.250
$\delta = 0.5$	0.076	0.138	0.194	0.245	0.293
			$E(P_1/F)$		
$\delta = 0.1$	0.963	0.976	0.983	0.988	0.992
$\delta = 0.2$	0.900	0.929	0.947	0.960	0.970
$\delta = 0.3$	0.826	0.869	0.898	0.921	0.939
$\delta = 0.4$	0.745	0.800	0.840	0.872	0.900
$\delta = 0.5$	0.658	0.724	0.774	0.816	0.854
			$\mathrm{Min}(P_1/F)$		
$\delta = 0.1$	0.900	0.900	0.900	0.900	0.900
$\delta = 0.2$	0.800	0.800	0.800	0.800	0.800
$\delta = 0.3$	0.700	0.700	0.700	0.700	0.700
$\delta = 0.4$	0.600	0.600	0.600	0.600	0.600
$\delta = 0.5$	0.500	0.500	0.500	0.500	0.500
			$\mathrm{Max}(P_1/F)$		
$\delta = 0.1$	1.530	1.281	1.178	1.120	1.083
$\delta = 0.2$	1.800	1.443	1.290	1.200	1.140
$\delta = 0.3$	1.963	1.544	1.360	1.252	1.178
$\delta = 0.4$	2.052	1.600	1.400	1.281	1.200
$\delta = 0.5$	2.081	1.618	1.413	1.291	1.207

(continued)

Table I—*Continued*

Panel C: Skewed Binomial: $\theta = 1/p$ with Prob p, $\theta = 0$ with Prob $(1-p)$; $\mathrm{Var}[\theta] = (1-p)/p$					
	$p = 0.1$	$p = 0.2$	$p = 0.3$	$p = 0.4$	$p = 0.5$
			$\lvert E(P_1–F)/F\rvert$		
$\delta = 0.1$	0.143	0.136	0.123	0.108	0.092
$\delta = 0.2$	0.260	0.249	0.227	0.200	0.170
$\delta = 0.3$	0.366	0.349	0.318	0.281	0.239
$\delta = 0.4$	0.465	0.440	0.400	0.352	0.300
$\delta = 0.5$	0.558	0.524	0.474	0.416	0.354

underreacted to F, or always overreacted, there would exist a profitable trading strategy that conditioned only on time-1 returns.

A more subtle result, and one that can be shown to hold generally—that is, for any distribution of θ—is that *on average*, the time-1 price underreacts to the fundamental F. The Appendix establishes the following:

PROPOSITION 1: *For any nondegenerate distribution of the random variable θ, there is underreaction on average, in the sense that $E(P_1/F) < 1$.*

The logic behind the proposition again flows from the zero-profit condition. In equilibrium, arbitrageurs behave as trend chasers, buying when returns are positive. Thus, their trading gains in those low-θ states when the price underreacts must be exactly offset by their losses in those high-θ states when the price overreacts. But note that in the high-θ states when there is overreaction—that is, when more arbs show up than expected—prices move more for a given fundamental shock and hence the linear return-based trading strategy buys a larger quantity of shares. So the way the zero-profit condition works out is that a smaller quantity times a larger expected profit on the underreaction side is balanced against a larger quantity times a smaller expected loss on the overreaction side.[6]

Finally, Table I shows that the average absolute distance between price and fundamentals, $E\,|(P_1 - F)/F\,|$, can be either greater or smaller with infinite arbitrage, as compared to the case of zero arbitrage. So the market can actually be less efficient in a price-equals-fundamentals sense when there is infinite arbitrage. This is more likely to be the case when the variance of θ is greater, as occurs in the skewed-binomial case shown in Panel C of the table. For example, if $\delta = 0.40$, while $\theta = 0$ with probability 0.8 and $\theta = 5$ with probability 0.2, we have that $E\,|(P_1 - F)/F\,| = 0.44$ with infinite arbitrage. In contrast, if there is no arbitrage whatsoever, $E\,|(P_1 - F)/F\,| = \delta = 0.40$.

[6] There is a winner's curse logic at work here that is reminiscent of Rock's (1986) model of IPO underpricing: Even though IPO prices are on average below first-day trading values, arbitrageurs who try to exploit this regularity by simply buying at the offer can nevertheless earn zero profits. This is because the allocation mechanism tends to give them more shares in those states of the world where the IPO price is too high, and fewer shares in those states of the world where it is too low.

To summarize, as arbitrage capacity grows large—that is, as N goes to infinity—expected returns to the simple trading strategy that exploits underreaction are driven to 0, implying that $\text{cov}(R_1, R_2) = 0$. Thus, increased competition among the arbitrageurs has the usual consequences for their trading profits. However, when there is uncertainty about the degree of crowding, the elimination of predictability does not correspond to prices being driven to fundamental values. In fact, the time-1 price does not converge in expectation to fundamentals. Even more strikingly, it is possible that prices can be further from fundamentals on average than in a world with no arbitrage at all.

One interpretation of these results is that with random θ, there are two distinct effects associated with increased values of N: (i) the familiar stabilizing impact of arbitrage; and (ii) a tendency for increased arbitrage capacity to also act as a form of endogenous noise trade, since when N is larger, there is more aggregate uncertainty about the total arbitrage response to any given event.[7] Note that the magnitude of the latter effect depends on the underreaction parameter δ: When δ is larger, arbitrageurs are induced to trade more aggressively. Therefore, to the extent that they impart noise to prices, this noise is magnified as δ increases. This is apparent in Table I. By contrast, holding δ fixed, arbitrageurs trade less aggressively when there is more dispersion in θ. As a result, an increase in the dispersion of θ increases the on-average underreaction effect described in Proposition 1; this too can be seen in Table I.

D. An Illustration: The MSCI Rebalancings of 2001–2002

The above ideas can be illustrated with a brief case study.[8] In December 2000, Morgan Stanley Capital International (MSCI) announced that it would be changing the methodology used to construct its MSCI indexes from a system of market-capitalization weighting to one of free-float weighting. This change was to be implemented in two phases. The first effective date (when the index would be partially revised) was set for November 30, 2001. The second effective date (when the index would be fully revised) was set for May 31, 2002. At the time, Lehman Brothers (2000) estimated that passive index managers worldwide would be forced to make trades totaling over $54 billion in order to conform to the new regime.

[7] By analogy to Grossman-Stiglitz (1980), one might say that an increase in N is like a simultaneous increase in both the number of rational uninformed traders (who do not observe fundamentals but who make inferences from prices) and the number of noise traders. Alternatively, when N goes to infinity and predictability in returns is eliminated, it must be that price equals the expected value of fundamentals conditional on *some* information set. Thus, arbitrageurs move the market from an equilibrium in which price is a biased reflection of the signal F to one in which price is an unbiased reflection of a degraded signal, namely, a noisy combination of F and θ; herein lies the central tradeoff. I thank John Campbell for suggesting the latter interpretation.

[8] I am grateful to Robin Greenwood for bringing this case to my attention, and for pointing me in the direction of the relevant research reports. Hau (2008) uses this case to explore a different set of asset pricing issues.

Because of tracking error considerations, index funds have an incentive to delay some fraction of their portfolio adjustment until the two effective dates. This creates an opportunity for arbitrageurs to buy stocks that are to be upweighted, and short stocks that are to be downweighted, between the announcement and the effective dates. In an idealized world, this arbitrage would bring the entire price impact forward to the announcement date, and there would be no further price movement on the effective dates.

However, if arbitrageurs trade in a price-insensitive manner, without conditioning their demands for each stock on an estimate of what they think its ultimate post-effective date value will be, an effect similar to the one in the model may arise.[9] For example, suppose that on the initial announcement, the price of a stock that is to be upweighted jumps from 100 to 108. One possibility is that this reflects early demand from index funds, and that there will be more of this demand on the effective date. In this case, the long-run price of the stock may settle at 112, and an arbitrageur who unconditionally buys all upweighted stocks will make a profit. Alternatively, it may be that much of the initial rise is due to buying by other arbitrageurs, and that when they unwind on the effective date, the long-run price of the stock may settle back at 104. In this case, it would be a mistake to buy at 108 because the trade is already overcrowded.

Something like this overcrowding phenomenon appears to have taken place on the first effective date, November 30, 2001. According to Lehman Brothers (2001), a strategy that was long stocks to be upweighted and short stocks to be downweighted lost 6.18% on this one day.[10] In a post-mortem, Lehman noted the "sharp contrast to past rebalances," concluding that "perhaps the dynamics of index rebalancing have changed."

Interestingly, however, things were reversed on the second effective date, May 31, 2002. On this day, Lehman Brothers (2002) reports that the conventional arbitrage strategy paid off handsomely, generating a 6.30% positive return: "The May MSCI reconstitution trade behaved as expected. The portfolios of stocks that were an expected net buy... by passive managers went up, and the expected sells went down."

Taken together, these two events highlight the central economic intuition of the model. On the one hand, the average return to the arbitrage strategy across the two MSCI effective dates was almost exactly zero, consistent with this being a well-publicized and heavily traded opportunity. On the other hand, it is questionable whether this zero-profit state of affairs could be said to reflect the workings of a textbook efficient market: There was a great

[9] By contrast, if the representative arbitrageur takes an active view as to the post-effective date value, and formulates his demand based on the spread between this estimate and the current price, there can be no overcrowding—if an unexpectedly large number of arbs shows up, the perceived spread narrows, and everybody's demand can be adjusted accordingly.

[10] More specifically, the Lehman Brothers strategy involved taking a long position in stocks with more than 10 days of "buying pressure" (that is, predicted demand from indexers in excess of 10 days worth of normal trading volume) and a short position in stocks with more than 10 days of "selling pressure."

deal of nonfundamental volatility in upweighted and downweighted stocks on the two effective dates, due in large part to the trading of the arbitrageurs themselves.

E. Dynamics: Learning about the Degree of Crowding

In an effort to keep things simple, the model focuses on a static situation where the distribution of arbitrage capacity n is exogenously specified. A potentially interesting extension might be to introduce a dynamic element, with n evolving stochastically and with arbs trying to learn about n by observing the past returns of their trading strategies. Such an approach might shed light on a number of issues, including the extent to which fluctuations in the profitability of a given strategy can be expected to persist over time.

The learning problem facing arbitrageurs in this setting is likely to be quite challenging. Suppose a strategy—say, buying low-accrual stocks—yields abnormally high returns over an interval from time t to $t + k$. How should this observation change one's estimate of the degree of crowding in the strategy? On the one hand, the high returns might suggest that there were relatively few arbs active at time t, so low-accrual stocks were cheap at this point. Alternatively, the high returns could reflect a spike of entry into the strategy *between* t and $t + k$, which pushed up the price of low-issuance stocks. So, without putting any further structure on the problem, it is not even clear whether one's estimate of n should be revised up or down based on the recent performance of the strategy.

F. Related Themes

The above model can be thought of as a particular rebuttal to Friedman's (1953) famous claim that rational arbitrage necessarily brings prices closer to fundamentals. In this sense, it is related to previous work by Hart and Kreps (1986), Stein (1987), DeLong et al. (1990b), and Hong and Stein (1999), among others. The connection to the latter paper is probably the closest, in that Hong and Stein also show how an "unanchored" arbitrage strategy—trading on momentum—can exploit a tendency for the market to underreact to news, while at the same time widening the average gap between prices and fundamentals. However, the mechanism here, having to do with uncertainty about aggregate arbitrage capacity, is fundamentally different from that in Hong and Stein, and potentially applies to a much wider range of real-world trading strategies.

The model also shares with Abreu and Brunnermeier (2002, 2003) an emphasis on coordination problems in arbitrage, with the common theme being that each individual arbitrageur's life is complicated by the fact that he cannot perfectly predict the actions of his peers. In Abreu and Brunnermeier, each arb is uncertain as to *when* the others will act; here, each is uncertain about *how many* others will act.

II. The Determinants of Arbitrageur Leverage

A. Overview

Both real-world events as well as a growing academic literature have shown that leveraged arbitrageurs can sometimes have a destabilizing impact on markets. The mechanism is now a familiar one: If an arbitrageur takes a highly leveraged position in a given trade, even a small adverse movement in prices can wipe out a large fraction of his capital. If borrowing capacity is limited by available capital (i.e., there is effectively a downpayment constraint), this reduction in capital will force the arbitrageur to liquidate some of his holdings, thereby further pushing down prices. To the extent that other arbitrageurs have taken on similar leveraged positions, their capital is also depleted, and they too are forced to liquidate. The end result can be a fire sale: correlated selling pressure and a contagious downward spiral in prices. This framework has been helpful in thinking about both dramatic market events, such as the collapse of LTCM in 1998 and the quant crisis of August 2007, as well as more subtle but pervasive phenomena, such as patterns of conditional skewness in currency returns.[11]

What this story leaves unanswered is why, given the dangers, arbitrageurs would choose to be highly levered in the first place. In a fire-sale scenario, there is a clear advantage to any arbitrageur who remains unlevered—he is not forced to sell at a temporarily depressed price. Indeed, if the arbitrageur is sufficiently conservative, and keeps some financial slack available in this state of the world, he can profit by *buying more* of the distressed asset. Thus, lurking in the background of the fire-sales story is a classic capital structure question: What is the optimal mix of borrowing and permanent capital for an arbitrageur who has unlimited access to both forms of financing? And given this optimal mix, will there still be fire sales in equilibrium?

Existing treatments have tended to set aside this capital structure question by simply assuming that arbitrage capital is small relative to the scale of available trading opportunities, meaning that arbs have no choice but to lever up. For example, Gromb and Vayanos (2002) write: "... by fixing the [capital] of the arbitrageurs, we do not allow for entry into the arbitrage industry, which seems a realistic assumption for understanding short-run market behavior." (p. 368) By contrast, I am interested in a longer-run question: Whether, over time, the inflow of new capital into the industry will tend to drive down arbitrageur leverage, and hence mitigate the associated problems. This long-run question brings the capital structure issue front-and-center.

To address this question, I proceed as follows. I assume that each arbitrageur can either raise permanent equity capital at a per-unit cost of c, or can borrow on a short-term basis at a zero interest rate. Thus, c represents the incremental

[11] See, for example, Brunnermeier, Nagel, and Pedersen (2008). They show that high-interest rate currencies tend to have more negatively skewed returns than low-interest rate currencies, and argue that this conditional skewness is a consequence of arbitrageur leverage, given that arbs have a general propensity to invest in "carry trades" that are long high-interest rate currencies, and short low-interest rate currencies.

cost of long-term equity finance, which could arise because equity investors—who are less protected than short-term lenders—need to expend resources on monitoring fund management. The leverage ratio is capped, in that an arbitrageur with one dollar of equity capital can take a total position of no more than L dollars, thereby borrowing at most $(L - 1)$ dollars.[12]

Given the wedge between the costs of debt and equity, there is the following basic trade-off. On the one hand, to the extent that prices are expected to converge smoothly to fundamentals (the "good" state), it is optimal for arbitrageurs to lever to the maximum to take advantage of cheap debt. On the other hand, if there is a nonzero probability of an adverse price move prior to convergence (the "bad" state), this may favor a more conservative capital structure. This is analogous to the standard corporate finance logic that debt is less attractive when there is a higher probability of financial distress.

What makes things more complicated in this setting, however, is that the terms of the trade-off depend critically on equilibrium prices in the two states, and these prices are in turn a function of c and of arbitrageur capital structure. Consider first the limiting case in which $c = 0$. If arbs choose to be unlevered, the infinite supply of costless equity capital implies that prices must be driven all the way to fundamental values—that is, the initial spread on any trade must be zero. But then it can never be profitable for any one arbitrageur to deviate to a higher-leverage strategy: With zero returns in the good state, there is nothing to lever up, and hence no point in choosing a riskier capital structure. Thus, when equity is literally costless, the long-run equilibrium involves all-equity financed arbitrageurs and there are never any fire sales. Similar logic applies when c is arbitrarily small but nonzero.

However, things change dramatically if the cost of equity c rises further. Now, if the arbs remain unlevered, returns in the good state must be sufficiently positive for them to recoup their cost of equity. Moreover, given that everybody is unlevered, adverse price movements in the bad state of the world are small. Taken together, these two factors can destroy the unlevered equilibrium: Any one arbitrageur may prefer to deviate to a policy of maximum leverage because this allows him to better exploit the small positive returns in the good state without paying too much of a price in the bad state (on the conjecture that everybody else remains unlevered). The bottom line is that for larger values of c, the only Nash equilibrium in the capital structure game is for all arbitrageurs to be maximally levered, even though this leads to large price crashes in the bad state.

B. Market Structure

I build on the same market structure as in Shleifer and Vishny (1997). There is an asset that will pay a final dividend of V at time 3. At time 1, there is

[12] The model of Acharya and Viswanathan (2008) can be thought of as endogenizing the maximum leverage ratio for an arbitrageur, based on moral hazard effects. See also Fostel and Geanakoplos (2008). A key difference relative to these papers is that I allow arbitrageurs in the aggregate to raise an arbitrarily large amount of equity financing—that is, external finance need not take the form of short-term debt.

a noise shock of δ_1, and arbitrageurs enter and try to offset this shock. If the total dollar value of the arbitrage position at time 1 is given by a_1, then the price is $P_1 = V - \delta_1 + a_1$.[13] At time 2, there is a probability p of a bad state in which the noise shock deepens to $\delta_2 > \delta_1$. In this case, the price becomes $P_{2d} = V - \delta_2 + a_2$, where a_2 is the total dollar value of the arbitrage position taken at time 2. Alternatively, there is a probability $(1 - p)$ of a good state in which all noise disappears by time 2, and the asset's price converges early to its fundamental value: $P_{2c} = V$. The good state is taken to be the more likely "normal" state of affairs, implying that p is small; this will be made more precise shortly.

There is a large pool of risk-neutral arbitrageurs. Each arb can raise one dollar of permanent equity capital, which requires the payment of an up-front cost of c at time 1. The notion that capital is permanent means that equity providers cannot withdraw their funds early, at time 2.[14] In addition, arbs can choose to lever this equity, by borrowing at time 1 from lenders who are also risk-neutral, and who demand a zero expected return. As noted above, the time-1 leverage ratio is bounded, so that an arb with one dollar of equity can borrow up to a maximum of $(L - 1)$ dollars. Importantly, all borrowing is short term: It must be repaid at time 2. This implies that in the bad state of the world, when the noise shock deepens, the arbs must settle their debts and can only invest out of any equity that remains.[15] There is also no scope for any further equity-raising at time 2.

C. Possible Arbitrage Strategies

For each dollar of equity capital, an arbitrageur can borrow anywhere between zero and $(L - 1)$ dollars. Hence, his investment at time 1, denoted by I, must satisfy $0 \leq I \leq L$. A first step in the analysis is to characterize the set of strategies that can be individually optimal for an arbitrageur. This is done in the following lemma. (The Appendix contains proofs of any lemmas or propositions not established in the text.)

LEMMA 1: *Only three types of strategies can be optimal for an individual arbitrageur: (i) a "max-leverage" strategy in which $I = L$, and in which the arbitrageur's wealth may be totally wiped out if the bad state occurs at time* 2; *(ii) a "cautious" strategy in which $0 < I < L$, and in which leverage is sufficiently modest that, given equilibrium prices, the arbitrageur's time-*2 *wealth is strictly*

[13] As in Shleifer and Vishny (1997), this reduced-form pricing rule can be thought of as reflecting the demands of an unmodelled third group of fundamental-based traders.

[14] This is in contrast to Shleifer and Vishny (1997), who focus on the performance-flow relationship and who therefore allow equity to flow out at time 2 in response to poor performance.

[15] An alternative approach that might seem more realistic is to allow arbitrageurs to roll over their borrowing at time 2, so long as they continue to satisfy the leverage ratio constraint. This yields results similar to those I derive below. This is because, in any equilibrium where arbitrageurs lever to the max at time 1, they end up losing all of their equity in the bad state at time 2. With no equity, and a finite leverage ratio constraint, their investment in this state is therefore zero either way.

positive in the bad state; and (iii) a "waiting" strategy in which $I = 0$, that is, the arb holds all of his equity back for time 2.

The intuition for the lemma is that with risk neutrality, the arbitrageur's problem is a linear one, so he can be driven to corner solutions. If good-state returns are attractive enough, he will invest as much as possible at time 1; this is the max-leverage strategy. Conversely, if the fire-sale discount in the bad state is deep enough, he will hold back all of his resources so he can be a buyer in that state; this is the waiting strategy. It is only when these two effects are just in balance that the interior cautious strategy becomes optimal, with the arb investing some at time 1 and some (in expectation) at time 2.

More precisely, let $\Delta_1 \equiv (V - P_1)/P_1$ represent the percentage return to investment at time 1, and let $\Delta_2 \equiv (V - P_{2d})/P_{2d}$ represent the return to investment at time 2, if the bad state occurs and the noise shock deepens. If an arb invests at time 1 and holds the position until convergence, he earns a return of Δ_1. If he has nonzero wealth to invest at time 2, he earns a return of Δ_2 with probability p on that wealth. This implies

LEMMA 2: *A necessary condition for the cautious strategy to be optimal is that equilibrium prices satisfy $\Delta_1 = p\Delta_2$. A necessary condition for the waiting strategy to be optimal is that equilibrium prices satisfy $\Delta_1 < p\Delta_2$.*

Lemma 1 describes the strategies that can potentially be optimal for an *individual arbitrageur*. However, if we want to focus on symmetric equilibria, in which all arbs pursue the same strategies, we can narrow things down further.

LEMMA 3: *Assume that $\delta_1/(V - \delta_1) > p\delta_2/(V - \delta_2)$. There can never be a symmetric equilibrium in which all arbs play the waiting strategy.*

The condition in Lemma 3 is easily satisfied for small values of p, and I assume that it holds in everything that follows. The idea behind the lemma is that if *all* arbs play the waiting strategy, there is no corrective force at time 1, so prices deviate significantly from fundamentals. But if this is the case, it cannot make sense to pass up the time-1 trading opportunity, especially if the likelihood of a time-2 opportunity ever arising is small (that is, if p is low).

The upshot of the analysis to this point is that if we want to focus on symmetric equilibria, we can restrict attention to two candidates: an equilibrium in which all arbs play the max-leverage strategy, and one in which they all play the cautious strategy. The waiting strategy is only observed in a mixed equilibrium in which some arbs play this strategy while at the same time others play the max-leverage strategy. I discuss this mixed equilibrium later, after first establishing the conditions under which either of the two symmetric equilibria can exist.[16]

[16] The waiting strategy also plays a role in an off-equilibrium-path sense, since the potential for individual arbs to deviate to this strategy can upset the max-leverage equilibrium for some parameter values. This will be seen explicitly below.

D. The Max-Leverage Equilibrium

Suppose that n arbitrageurs enter the market, and all play the max-leverage strategy. In this case, the time-1 price is given by $P_1 = V - \delta_1 + nL$. If the bad state hits at time 2, the equity position of each arb is $w_2 = 1 + L(P_{2d}/P_1 - 1)$. This expression incorporates the fact that an arb must pay back his loan at time 2. Note that as calculated, equity can be negative, which should be interpreted as a situation where the arb is wiped out (his equity is actually zero) and unable to fully repay his debts.

It follows that a maximally leveraged arbitrageur is wiped out if $L > P_1/(P_1 - P_{2d})$. Clearly, it is easier to satisfy this condition if the upper bound L on leverage is relatively high. As it turns out, the analysis is considerably simplified—with no real loss of insight—if we focus attention on the (high-L) subset of the parameter space where this wipe-out condition is satisfied. I therefore proceed on the premise that it is in fact satisfied, and then, after having solved for the endogenous variables P_{2d} and P_1, will come back and rewrite the condition in terms of the primitive parameters of the model.

If all arbitrageurs are fully wiped out in the bad state at time 2, there is no offset at all to the noise shock, so the price is given by

$$P_{2d} = V - \delta_2. \tag{10}$$

To pin down the time-1 price, we need to analyze the entry decisions of the arbitrageurs. What matters here are the returns to the arbitrageur-lender *coalition*. In other words, the equilibrium entry condition is that the expected total dollar return to a leveraged position just equals the upfront cost of equity c. This allows each active arbitrageur to pay off his debts in expectation, while also covering the cost of equity.[17]

Not including the equity cost c, the expected return to a position of size L is given by $(1-p)L\Delta_1 + pL(P_{2d} - P_1)/P_1$. This reflects the fact that a leveraged position earns the return Δ_1 in the good state when there is no deepening of the noise shock, but has to be fully liquidated at a price of P_{2d} in the bad state, when there is such a deepening. It follows that the equilibrium entry condition is given by $(1-p)L\Delta_1 + pL(P_{2d} - P_1)/P_1 = c$. Using the expression for P_{2d} in (10), this can be manipulated to yield

$$P_1 = \frac{V - p\delta_2}{(1 + c/L)}. \tag{11}$$

Thus, in the postulated max-leverage equilibrium, prices are given by (10) and (11).[18] With these prices in hand, the above wipe-out condition that

[17] Implicit in all of this is that a lender has to charge a nonzero default premium so that lending profits in the good state offset defaults in the bad state. However, there is no need to compute this premium explicitly—all that we need to impose for the current purposes is that each arbitrageur and his lenders jointly break even in expectation.

[18] If one is interested in an explicit formula for n, the number of arbs who are active in equilibrium, this is obtained by setting the right-hand side of (11) equal to $V - \delta_1 + nL$.

$L > P_1/(P_1 - P_{2d})$ can be re-expressed in terms of primitive parameters as

$$L > \frac{V(1+c) - \delta_2(p+c)}{\delta_2(1-p)}. \tag{12}$$

This condition is easily satisfied for reasonable values of L. For example, with $V = 100$, $\delta_2 = 20$, $p = 0.05$, and $c = 0.1$, the condition is satisfied so long as $L > 5.63$.

It should be stressed that (10) and (11) describe a *candidate* equilibrium, on the assumption that all arbitrageurs play the max-leverage strategy. However, in order for the equilibrium to actually be sustainable, it must be the case that no arbitrageur prefers to deviate to an alternative strategy.

That such deviations are relevant when c is low can be seen by considering the limit when c goes to zero. In this case, we have that $P_1 = V - p\delta_2$ and $P_{2d} = V - \delta_2$. Thus, in spite of the fact that there is costless entry into arbitrage, the candidate max-leverage equilibrium involves prices diverging from fundamentals at both time 1 and 2. With these prices, levered arbs just break even, making a small return in the high-probability good state and being forced to liquidate at a large loss in the bad state.

If arbitrageurs were *restricted* to using max-leverage strategies, this would be a stable situation. But if arbs are free to choose alternative strategies, it cannot be. Given a zero cost of equity and a large discount in the time-2 bad state, it is obviously profitable to deviate to the waiting strategy, since one dollar invested in that strategy earns a strictly positive expected return of $p\delta_2/(V - \delta_2)$.

More generally, for any $c > 0$, this deviation to the waiting strategy remains an option for an individual arbitrageur. In order for the deviation to be unprofitable, it must be the case that the expected return to the waiting strategy is less than c. Thus, we have

PROPOSITION 2: *In order to sustain a symmetric equilibrium in which all arbitrageurs play the max-leverage strategy, it is necessary that* $c > c^+ = p\delta_2/(V - \delta_2)$.

The intuition behind the proposition is that if equity capital is very cheap, and everybody else is levered, it will make sense for somebody to deviate by building a war chest and waiting for a fire-sale buying opportunity—even if the probability of this opportunity arising is low. However, once the cost of equity rises a bit, waiting for a rare opportunity is no longer viable and the max-leverage equilibrium can be sustained.

E. The Cautious Equilibrium

In the cautious equilibrium, n arbitrageurs enter the market, and each invests an amount $I < L$. Moreover, I is sufficiently small that time-2 wealth in the bad state is strictly positive, that is, $w_2 = 1 + I(P_{2d}/P_1 - 1) > 0$, or alternatively, $I < P_1/(P_1 - P_{2d})$. It follows that prices are given by $P_1 = V - \delta_1 + nI$ and $P_{2d} = V - \delta_2 + nw_2$.

To solve for these prices in terms of exogenous parameters, we can use two conditions.[19] The first is Lemma 2, which says that in order for it to be optimal for arbs to play the cautious strategy, it must be that $\Delta_1 = p\Delta_2$. The second is an entry condition: The expected returns to the strategy must just equal the upfront cost of equity, implying that $\Delta_1 = p\Delta_2 = c$. Straightforward calculation based on these conditions yields

$$P_1 = V/(1+c), \tag{13}$$

$$P_{2d} = V/(1+c/p). \tag{14}$$

Note that in the limit when c goes to zero, we now have $P_1 = P_{2d} = V$. In other words, with an infinite supply of costless capital, cautious arbs drive prices to fundamentals in both periods. This is in contrast to the case with maximally levered arbs, where the zero-profit condition involves small positive returns in the good state and large fire-sale losses in the bad state. In this sense, the cautious strategy can be said to be more conducive to market efficiency.

Again, however, we must check whether the candidate equilibrium prices in (13) and (14) actually represent a stable equilibrium that is robust to potential deviations. Here the relevant deviation to be considered is that an arbitrageur might switch to playing the max-leverage strategy. If he does so, he earns a return (for the coalition of himself and his lenders) of $(1-p)L\Delta_1 + pL(P_{2d} - P_1)/P_1$. In order to sustain an equilibrium with cautious investors, this deviation has to be unprofitable, that is, the return must be less than the upfront equity cost c. Thus, we require $(1-p)L\Delta_1 + pL(P_{2d} - P_1)/P_1 < c$.

Plugging in the expressions for prices from (13) and (14), this requirement can be re-arranged to yield the following.

PROPOSITION 3: *In order to sustain a symmetric equilibrium in which all arbitrageurs play the cautious strategy, it is necessary that* $c < c^- = p/((1-p)L - 1)$.

Proposition 3 highlights the three forces that can undo the cautious equilibrium. First, as c rises above zero, the time-1 price falls below fundamentals and Δ_1 becomes positive—that is, a positive spread opens up. As the spread widens, so does the appeal of taking a highly leveraged position. This appeal is stronger when L is high, meaning that leverage can be used more aggressively, and when the probability p of being forced to unwind a leveraged position in distress is smaller.

The cautious strategy admits two distinct sub-cases. In the first, which might be termed a "modest-leverage strategy," $I > 1$. That is, arbitrageurs borrow, but not so much that they are ever in danger of being completely wiped out. In the second, which might be termed a "dry-powder strategy," $I \leq 1$. In this case, arbs remain completely unlevered and hold some of their equity capital back for

[19] As before, once one has solved for prices in terms of exogenous parameters, it is straightforward to go back and calculate what this implies for the equilibrium values of n and I.

time 2. It can be shown that the dry-powder strategy obtains when c is sufficiently small.

PROPOSITION 4: *In a cautious equilibrium, arbitrageurs play a dry-powder strategy with $I \leq 1$ if $c \leq c^{dp} = p(\delta_2 - \delta_1)/((V - \delta_2) - p(V - \delta_1))$.*

Thus, when the cost of equity c is small enough, arbitrageurs optimally choose to be very conservative. They rely entirely on equity finance, and keep a cash buffer on hand in case a fire-sale opportunity arises at time 2. This makes it clear why, in the limit when c goes to zero, it must be the case that $P_1 = P_{2d} = V$: There is always enough spare arbitrage capacity held back at time 2 to completely offset any deepening of the noise shock.

F. A Mixed Equilibrium

It is straightforward to verify that if the wipe-out condition in (12) holds (that is, L is sufficiently high), then $c^- < c^+$. Thus, the max-leverage and the cautious equilibrium regions never overlap. Instead, there is an intermediate region, for values of c such that $c^- \leq c \leq c^+$, where neither of these symmetric equilibria are viable. In this region, the only equilibrium that can be sustained is a mixed equilibrium in which two different arbitrage strategies co-exist.

In the mixed equilibrium, n_L arbitrageurs play the max-leverage strategy and n_W play the waiting strategy. It follows that prices are given by $P_1 = V - \delta_1 + Ln_L$ and $P_{2d} = V - \delta_2 + n_W$.[20] As before, the key to expressing these prices in terms of exogenous parameters is to make use of zero-profit conditions. For the arbs who play the waiting strategy, and who therefore have a probability p of being able to earn the return Δ_2, the breakeven condition is that $p\Delta_2 = c$. It follows that the time-2 price is still given by equation (14) above, that is, it is the same as in the cautious equilibrium.

For the max-levered arbs, the breakeven condition continues to be that $(1 - p)L\Delta_1 + pL(P_{2d} - P_1)/P_1 = c$. The only difference is that the time-2 price is now given by (14), instead of by (10) as in the max-leverage equilibrium. Making this substitution, we have

$$P_1 = \frac{V(1 - cp/(c + p))}{(1 + c/L)}. \tag{15}$$

By construction, both the waiting strategy and the max-leverage strategy earn zero profits (net of the upfront cost of c) in the mixed equilibrium, so there is no temptation for an arb playing either strategy to deviate to the other. It is also straightforward to check that given prices in the mixed equilibrium, we have that $\Delta_1 < p\Delta_2$, so long as $c \geq c^-$. From Lemma 2, this implies that the waiting strategy dominates the cautious strategy. Hence, there are no deviations that overturn the mixed equilibrium within the region of interest.

[20] The expression for P_{2d} presumes that max-levered arbitrageurs are wiped out in the bad state at time 2 and hence contribute nothing to the price. This condition can be shown to hold whenever $c \geq c^-$.

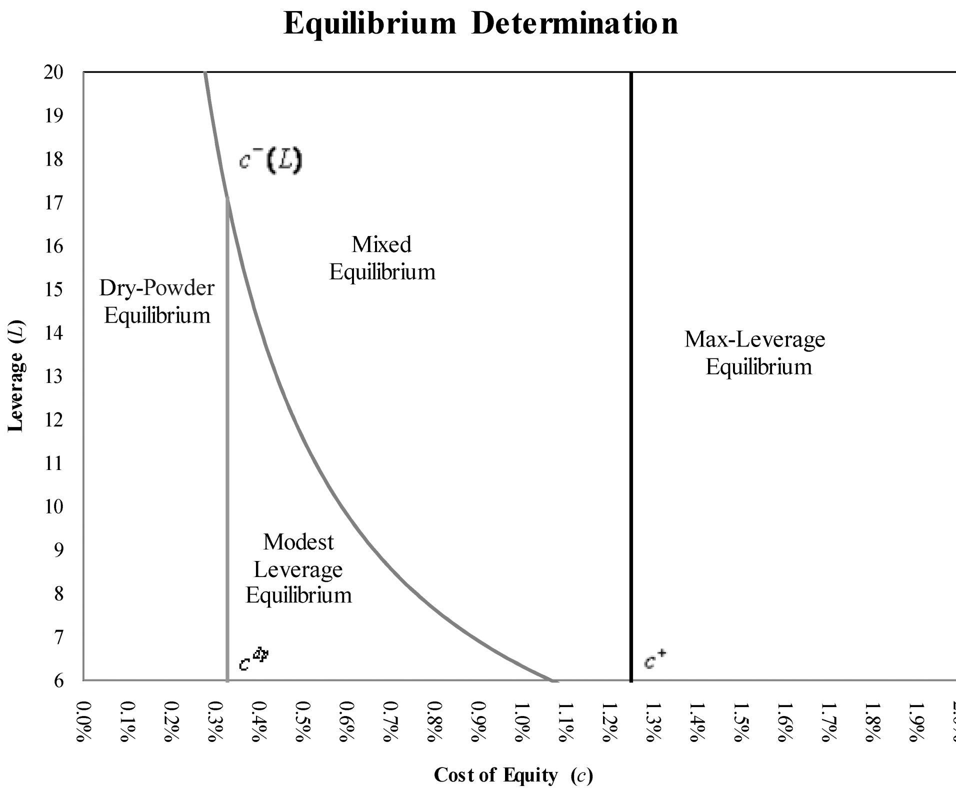

Figure 1. Equilibria of the leverage model. The figure illustrates the division of the parameter space into the different equilibrium regions, setting $V = 100$, $\delta_1 = 15$, $\delta_2 = 20$, and $p = 0.05$, and allowing the equity cost c to vary between zero and 0.02, while the maximum leverage L varies between 6 and 20.

The following proposition summarizes the analysis to this point.

PROPOSITION 5: *Assume that (12) holds. If $c < c^- = p/((1-p)L - 1)$, the cautious equilibrium obtains, with $P_1 = V/(1+c)$ and $P_{2d} = V/(1+c/p)$. Within the cautious region, arbs play a dry-powder strategy if $c \leq c^{dp} = p(\delta_2 - \delta_1)/((V-\delta_2) - p(V-\delta_1))$ and a modest-leverage strategy otherwise. If $c > c^+ = p\delta_2/(V-\delta_2)$, the max-leverage equilibrium obtains, with $P_1 = (V - p\delta_2)/(1+c/L)$ and $P_{2d} = V - \delta_2$. If $c^- \leq c \leq c^+$, the mixed equilibrium obtains, with $P_1 = (V(1 - cp/(c+p)))/(1+c/L)$ and $P_{2d} = V/(1+c/p)$. Both P_1 and P_{2d} are continuous functions over the entire parameter space, including the boundaries defining the different regions.*

Figure 1 illustrates the division of the parameter space into the different regions, setting $V = 100$, $\delta_1 = 15$, $\delta_2 = 20$, and $p = 0.05$, and allowing c to vary between zero and 0.02, while L varies between 6 and 20.

G. Pros and Cons of Leverage Limits

One way to gain further insight into the model's implications is to consider the consequences of a regulation that constrains arbitrageur leverage. In particular, suppose the private market allows arbs to lever themselves up, but that a regulator imposes a complete ban on borrowing, thereby effectively setting $L = 1$. How do prices compare in the regulated and unregulated cases?

A first observation is that regulation is nonbinding if $c \leq c^{dp}$, since in this case arbitrageurs optimally choose to forgo leverage, even if they are allowed to use it. So prices in the regulated and unregulated cases coincide and continue to be given by the cautious-equilibrium values in (13) and (14). By contrast, when $c > c^{dp}$, the regulation binds. This implies that regulated arbs invest as much as they can at time 1, that is, that $I = 1$. It is straightforward to show (see Appendix) that in this case, the time-1 price is again given by (13), but that the time-2 price is modified to

$$P_{2d}(\text{regulated}) = \frac{V(V - \delta_2)}{(1 + c)(V - \delta_1)}. \tag{16}$$

Figure 2 plots prices as a function of c in both the regulated and unregulated cases, for an example in which, as above, $V = 100$, $\delta_1 = 15$, $\delta_2 = 20$, and $p = 0.05$, and in which $L = 10$ when there is no regulation. For these parameter values, $c^{dp} = 0.0033$, $c^{-} = 0.0059$, and $c^{+} = 0.0125$. The figure points to three general conclusions. First, as just noted, regulation is irrelevant for the lowest values of c, namely, when $c \leq c^{dp}$.

Second, regulation looks attractive for intermediate values of c. Time-2 prices in the bad state are substantially higher under regulation because crashes are avoided. And as long as c is not too large, time-1 prices are very close in the regulated and unregulated cases; indeed, they are identical if $c < c^{-}$. To be specific, consider what happens when $c = 0.0125$, so that we are just inside the max-leverage region. Time-1 prices are 98.88 and 98.77 in the unregulated and regulated cases, respectively—almost the same. However, time-2 prices in the bad state are 80.00 and 92.96—that is, much higher in the regulated case. Thus, regulation effectively gets rid of leverage-induced crashes, at almost no cost in terms of distorting time-1 prices.

Third, for higher values of c, regulation is more of a double-edged sword. It still has the beneficial effect of preventing crashes at time 2, but now it meaningfully impedes arbitrage at time 1. For example, with $c = 0.050$, time-1 prices are 98.51 and 95.24 in the unregulated and regulated cases, respectively. Intuitively, when equity financing becomes very expensive, a ban on leverage can significantly reduce total arbitrage activity at time 1, leading to a wider spread between prices and fundamentals.

Of course, the model is too stylized to yield decisive conclusions about the merits of imposing leverage limits on arbitrageurs. Nevertheless, it does highlight a key point: When there are fire-sale externalities associated with leverage

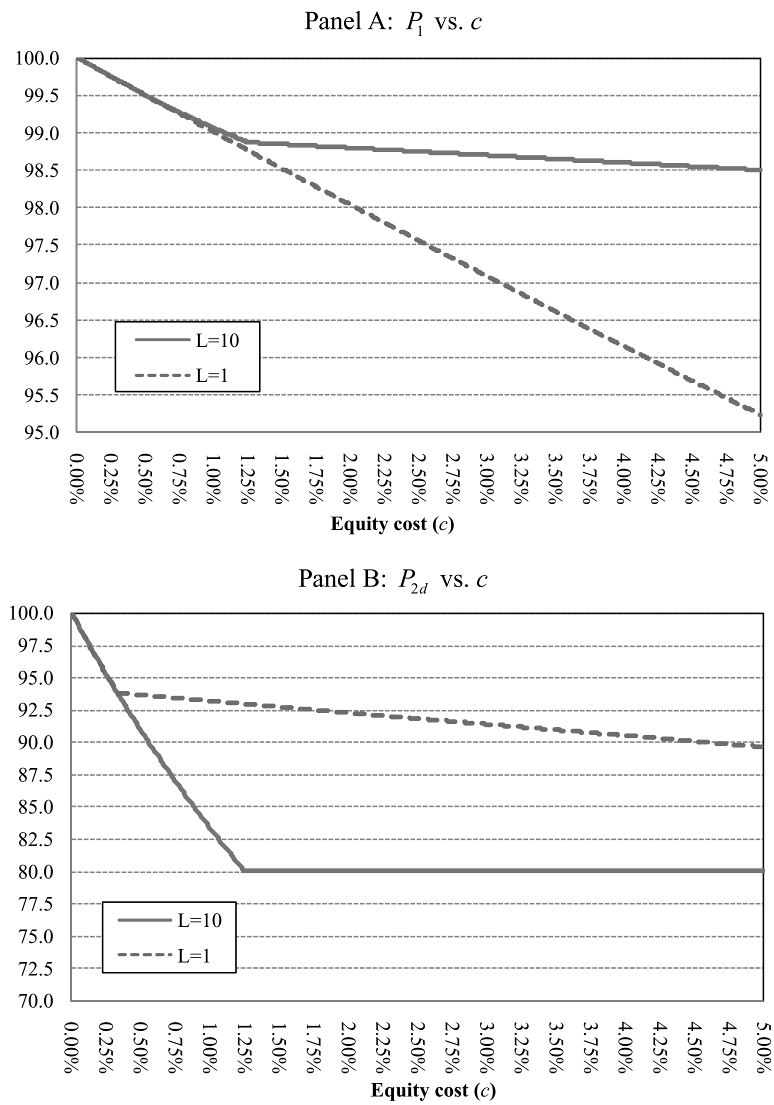

Figure 2. Prices as a function of c, with and without a ban on leverage. The figure plots the time-1 price P_1 and the bad state time-2 price P_{2d} against the equity cost c, in both the unregulated case (where maximum leverage $L = 10$) and the regulated case (where $L = 1$). The other parameters of the model are set as follows: $V = 100$, $\delta_1 = 15$, $\delta_2 = 20$, and $p = 0.05$.

choice, there is no general presumption that individually rational decisions on the part of arbs will lead to the most efficient configuration of prices.[21] This point emerges most starkly in the intermediate c case: Prices are much closer to fundamentals when leverage is constrained by regulation than when arbs are allowed to choose leverage freely.

[21] Stein (2005) makes a related point about arbitrageurs' choice of whether to operate in an open-end versus closed-end form: Individually rational arbs may choose to be open-ended, even though

H. Broader Implications for Financial Regulation

At a more general level, the model developed above suggests a few basic messages for financial regulation. First, it offers a nontraditional way of thinking about capital requirements. The usual rationale for capital requirements, which is focused on the banking industry, begins with the observation that government deposit insurance gives banks an incentive to take undue levels of risk. Mandated capital levels are meant to temper such risk-shifting incentives, and to protect the deposit insurer from any resulting losses. By contrast, in the current model there is no insurance and private lenders do a fine job of looking out for their own interests—they break even in expectation. Thus, the goal of any capital requirements regime would not be to protect senior claimants per se, but rather to mitigate fire-sale spillovers across institutions.

This difference in perspective leads to different implications for how capital requirements should be designed. For example, in the Basel-II framework, the capital requirement for a bank that holds a particular asset depends on the risk attributes of that asset *taken in isolation*. However, if the focus is on reducing fire-sale spillovers, one would ideally want the capital requirement to depend on: (i) the likelihood that the asset would be sold in a distress situation and (ii) the extent to which other highly leveraged entities are holding the same asset.[22]

A second message is that in some cases capital regulation can be quite distortionary itself. In the model, this occurs when c is relatively high and a cap on leverage significantly reduces aggregate arbitrage activity. In such cases, an important question to ask is whether some form of ex post intervention in extreme states can be more efficient than trying to solve everything with very stringent ex ante capital requirements. Such ex post intervention might include the government acting as a market-maker of last resort, buying up assets that are subject to overwhelming systemic fire-selling—as in the first incarnation of the Treasury's Troubled Assets Relief Program (TARP). Or it might involve the government making financing available on favorable terms to private liquidity providers who provide a similar function.

A final message is that policymakers might want to think more about ways to narrow the wedge between the perceived costs of short-term debt and equity finance.[23] As the model makes clear, the leverage choices of individual arbitrageurs are highly sensitive to c in equilibrium, and when c is low enough, arbs voluntarily choose to finance themselves in such a conservative way that

the market as a whole might be more efficient if it were populated with closed-end funds. Caballero and Krishnamurthy (2003) note that, in emerging markets, private firms' decisions about which currency to borrow in may also not be socially optimal—there is a tendency toward excessively risky dollar-denominated debt.

[22] Adrian and Brunnermeier (2008) develop a methodology aimed at measuring such spillover risks.

[23] This point is emphasized by Kashyap, Rajan, and Stein (2008). Diamond and Rajan (2001) and Diamond (2004) argue that short-term debt is particularly attractive to financial firms because of its role in containing agency problems.

no regulation is needed. Getting a better handle on the underlying frictions that determine c in financial firms, and on how these frictions might be reduced, seems like an important task for ongoing research.

III. Conclusions

It is undeniable that sophisticated professional investors play a more dominant role in financial markets than they used to. A more difficult question is whether this form of progress will ultimately help to make markets more efficient. The analysis of this paper suggests an ambiguous answer: Because of the externalities associated with crowding and leverage, there is no clear theoretical presumption that—absent any policy intervention—the rise of sophisticated investors should necessarily be beneficial to market efficiency, even over a very long horizon. It will be interesting to see what the future brings.

Appendix

A. The Crowded-Trade Problem

Proof of Proposition 1: A generalization of the Law of Total Variance implies that $0 = \text{cov}[R_1, R_2] = \text{cov}[E[R_1 \mid F], E[R_2 \mid F]] + E[\text{cov}[R_1, R_2 \mid F]]$, or

$$0 = E\left[\frac{(1-\delta)(\delta-\theta\Phi^*)}{(1-\theta\Phi^*)^2}\right] = E\left[\frac{1-\delta}{1-\theta\Phi^*}\right]E\left[\frac{\delta-\theta\Phi^*}{1-\theta\Phi^*}\right] + \text{Cov}\left[\frac{1-\delta}{1-\theta\Phi^*}, \frac{\delta-\theta\Phi^*}{1-\theta\Phi^*}\right]. \tag{A1}$$

Because $(1-\delta)/(1-\theta\Phi^*)$ is increasing in θ and $(\delta-\theta\Phi^*)/(1-\theta\Phi^*)$ is decreasing in θ, we have that

$$E[\text{cov}[R_1, R_2 \mid F]] = \text{cov}[(1-\delta)/(1-\theta\Phi^*), (\delta-\theta\Phi^*)/(1-\theta\Phi^*)] < 0.$$

To see the intuition, note that once we condition on F, returns reflect random variation in the number of arbs. Conditional on $F > 0$, high (low) entry generates high (low) returns at time 1 and low (high) returns at time 2, and vice versa for $F < 0$. This implies that

$$\text{cov}[E[R_1|F], E[R_2|F]] = E[(1-\delta)/(1-\theta\Phi^*)] \times E[(\delta-\theta\Phi^*)/(1-\theta\Phi^*)] > 0. \tag{A2}$$

Next, note that the total price reaction at times 1 and 2, scaled by F, must sum to one for all realizations of θ. That is, $(1-\delta)/(1-\theta\Phi^*) + (\delta-\theta\Phi^*)/(1-\theta\Phi^*) = 1$. Taking expectations of this identity and substituting the result into (A2), we obtain

$$E[P_1/F](1 - E[P_1/F]) = E[(1-\delta)/(1-\theta\Phi^*)] \times (1 - E[(1-\delta)/(1-\theta\Phi^*)]) > 0, \tag{A3}$$

which implies that $0 < E[P_1/F] < 1$. Q.E.D.

Computations Underlying Table I: If one specifies the distribution of the random variable θ, as well as a value of δ, equation (9) can be solved for the equilibrium trading intensity Φ^*. If $\theta_L > 0$, it is clear that there is at least one solution to (9) on $(0, \delta/\theta_L)$. Although (9) may have more than one root, the relevant solution must satisfy $\Phi^* < 1/\theta_H$, which ensures that $P_1/F > 0$ for all realizations of θ.

1. *Symmetric Binomial*: Assume that $\theta = (1-h)$ or $\theta = (1+h)$, each with probability $\frac{1}{2}$. In this case, Φ^* is the smallest positive root of a cubic equation

$$0 = (h^2-1)\Phi^3 + (\delta + \delta h^2 - 2h^2 + 2)\Phi^2 - (1+2\delta)\Phi + \delta. \tag{A4}$$

It can be shown that (A4) has a root satisfying $\Phi^* < 1/(1+h)$.

2. *Uniform*: Assume that θ is uniformly distributed on $[1-h, 1+h]$. Here, Φ^* satisfies

$$0 = \frac{1-\delta}{(h^2-1)\Phi^2 + 2\Phi - 1} + \frac{1}{2h\Phi} \ln\left[\frac{(1-h)\Phi - 1}{(1+h)\Phi - 1}\right]. \tag{A5}$$

As above, one can verify that (A5) has a solution satisfying $\Phi^* < 1/(1+h)$.

3. *Skewed Binomial*: Assume that $\theta = 1/p$ with probability p, and $\theta = 0$ with probability $(1-p)$. In this case, Φ^* is the root of a quadratic polynomial, and the relevant solution, satisfying $\Phi^* < p$, is given by

$$\Phi^* = p\left[\left(\frac{p}{2\delta(1-p)-1} + 1\right) - \sqrt{\left(\frac{p}{2\delta(1-p)-1} + 1\right)^2 - \frac{1}{1-p}}\right]. \tag{A6}$$

B. The Determinants of Arbitrageur Leverage

I first lay out the arb's payoff function in some detail. This payoff functions turns out to depend crucially on whether or not the arb is wiped out in the bad state at time 2. Consider the payoff to an arb who makes a time-1 investment, I, satisfying $0 \le I \le L$. After repaying his loan, the time-2 wealth of the arb is $w_2(P_2) = 1 + I(P_2/P_1 - 1)$. Therefore, the arb is wiped out in the bad state if $w_2(P_{2d}) < 0$, or

$$I > I_W \equiv P_1/(P_1 - P_{2d}). \tag{B1}$$

The payoff function of the arbitrageur-lender coalition then takes the form

$$\Pi(I) = \mathbf{1}\{I > I_W\} \cdot \Pi_W(I) + \mathbf{1}\{I \le I_W\} \cdot \Pi_{NW}(I), \tag{B2}$$

where

$$\Pi_W(I) = I[(1-p)\Delta_1 + p(P_{2d}/P_1 - 1)] - c \tag{B3}$$

is the profit function if the arb is wiped out in the bad state at time 2 and

$$\Pi_{NW}(I) = [I\Delta_1 + p(1-I)\Delta_2] - c \tag{B4}$$

is the profit function if the arb is not wiped out. Thus, $\Pi(I)$ is a piecewise linear function of I with a kink at I_W.

To understand (B3), suppose the arb is wiped out in the bad state. With probability p, the bad state obtains, the portfolio is liquidated, and the return to the coalition is $I(P_{2d}/P_1 - 1)$. With probability $(1-p)$, the good state obtains and the return to the coalition is $I\Delta_1$. Thus, if $I > I_W$, the expected return to the coalition net of the equity cost is given by $\Pi_W(I)$.

To understand (B4), suppose the arb is not wiped out in the bad state. Now when the bad state obtains, investors have $w_2(P_{2d}) \geq 0$ to be re-invested at time 2. In this case, the coalition earns $(V/P_{2d})w_2(P_{2d}) - 1 = I\Delta_1 + (1-I)\Delta_2$. Thus, when $I \leq I_W$, the net expected return to the coalition is given by $\Pi_{NW}(I)$.

Proof of Lemmas 1 and 2: First, suppose prices are such that $I_W > L$, so arbs cannot be wiped out for any $I \in [0, L]$. If $\Delta_1 > p\Delta_2$, we are at a corner where $I^* = L$ and arbs play a max-leverage strategy. If $\Delta_1 = p\Delta_2$, arbs are indifferent between any $I \in [0, L]$. This corresponds to the cautious strategy described in the text, given that $I_W > L$ and arbs cannot be wiped out. Finally, if $\Delta_1 < p\Delta_2$, $I^* = 0$ and we are at the other corner, with arbs playing a waiting strategy.

Next, suppose prices are such that $0 < I_W \leq L$, so that it is possible for arbs to be wiped out if they invest enough at time 1. Let $\pi_{NW} = \Delta_1 - p\Delta_2$ denote the slope of $\Pi(I)$ for $I \leq I_W$ and $\pi_W = [(1-p)\Delta_1 + p(P_{2d}/P_1 - 1)]$ denote the slope for $I > I_W$. Since $P_{2d} < P_1$, we have that $\pi_{NW} < \pi_W$. It follows that the only possible outcomes are $I^* = 0$ (which requires $\pi_{NW} < 0$) or $I^* = L$. To see that no other value of I can be optimal in this region, consider what happens for the following permutations of π_{NW} and π_W:

	π_{NW}	π_W	Outcome
(1)	> 0	> 0	$I^* = L$
(2)	> 0	$= 0$	This case is ruled out because $\pi_{NW} < \pi_W$.
(3)	> 0	< 0	This case is ruled out because $\pi_{NW} < \pi_W$.
(4)	$= 0$	> 0	$I^* = L$
(5)	$= 0$	$= 0$	This case is ruled out because $\pi_{NW} < \pi_W$.
(6)	$= 0$	< 0	This case is ruled out because $\pi_{NW} < \pi_W$.
(7)	< 0	> 0	$I^* = 0$ if $\Pi(0) > \Pi(L)$, and $I^* = L$ if $\Pi(0) < \Pi(L)$.
(8)	< 0	$= 0$	$I^* = 0$
(9)	< 0	< 0	$I^* = 0$

Therefore, if $0 < I_W \leq L$ arbs play either a max-leverage strategy or a waiting strategy. The cautious strategy can never be optimal in this region. Q.E.D.

Proof of Lemma 3: Suppose that $n > 0$ arbitrageurs enter, all choosing $I = 0$. In this case, we have $P_1 = V - \delta_1$ and $P_{2d} = V - \delta_2 + n$, which implies that $\Delta_1 = \delta_1/(V - \delta_1)$ and $\Delta_2 = (\delta_2 - n)/(V + n - \delta_2)$. By Lemma 1, $I^* = 0$ can only be optimal if $\Delta_1 < p\Delta_2$, or $\delta_1/(V - \delta_1) < p(\delta_2 - n)/(V + n - \delta_2)$. However, if as assumed in the statement of Lemma 3, we have $\delta_1/(V - \delta_1) > p\delta_2/(V - \delta_2)$, this cannot be satisfied. Thus, there cannot be a symmetric equilibrium where all arbs play the waiting strategy. Q.E.D.

Proof of Proposition 4: Given that arbs are not wiped out in the cautious case, we can write $P_1 = V - \delta_1 + nI$ and $P_{2d} = V - \delta_2 + nw_2(P_{2d})$. We have also seen that equilibrium prices must satisfy $P_1 = V/(1 + c)$ and $P_{2d} = V/(1 + c/p)$, as in equations (13) and (14) in the text. Together, these facts allow us to solve for the equilibrium values of n and I. In particular, we have that

$$I^* = \frac{\delta_1 - cV/(1+c)}{[(1+c)(V-\delta_1)/(1+c/p) - (V-\delta_2)] + \delta_1 - cV/(1+c)}. \tag{B5}$$

This corresponds to a dry-powder strategy when $I^* < 1$. Using the above formula, it follows that $I^* < 1$ when

$$c \le c^{dp} = \frac{p(\delta_2 - \delta_1)}{(V-\delta_2) - p(V-\delta_1)}. \tag{B6}$$

Q.E.D.

Prices with Leverage Limits: Suppose a regulator fixes $L = 1$. For $c \le c^{dp}$, unregulated arbs would play a dry-powder strategy, so the constraint does not bind and prices are given by $P_1 = V/(1 + c)$ and $P_{2d} = V/(1 + c/p)$. For $c > c^{dp}$, the constraint binds and n arbs enter, each setting $I = 1$. It follows that $P_1 = V - \delta_1 + n$ and that $P_{2d} = P_1[(V - \delta_2)/(V - \delta_1)]$. The zero-profit condition is $0 = \Pi_{NW}(1) = \Delta_1 - c$, which implies $P_1(\text{regulated}) = V/(1 + c)$ and $P_{2d}(\text{regulated}) = [V/(1 + c)][(V - \delta_2)/(V - \delta_1)]$. These expressions assume that $c < \delta_1/(V - \delta_1)$: When leverage is not permitted, no arbs will enter if the upfront cost of equity is sufficiently high (that is, if $c > \delta_1/(V - \delta_1)$).

REFERENCES

Abreu, Dilip, and Markus K. Brunnermeier, 2002, Synchronization risk and delayed arbitrage, *Journal of Financial Economics* 66, 341–360.

Abreu, Dilip, and Markus K. Brunnermeier, 2003, Bubbles and crashes, *Econometrica* 71, 173–204.

Acharya, Viral V., and S. Viswanathan, 2008, Leverage, moral hazard and liquidity, Working paper, London Business School.

Adrian, Tobias, and Markus K. Brunnermeier, 2008, CoVaR, Working paper, Princeton.

Allen, Franklin, and Douglas Gale, 2005, From cash-in-the-market pricing to financial fragility, *Journal of the European Economic Association* 3, 535–546.

Asness, Clifford S., Tobias J. Moskowitz, and Lasse H. Pedersen, 2008, Value and momentum everywhere, Nyu Stern School Working Paper.

Bernard, Victor L., and J. Thomas, 1989, Post-earnings announcement drift: Delayed price response or risk premium? *Journal of Accounting Research* 27, 1–48.

Bernard, Victor L., and J. Thomas, 1990, Evidence that stock prices do not fully reflect the implications of current earnings for future earnings, *Journal of Accounting and Economics* 13, 305–340.

Brunnermeier, Markus K., Stefan Nagel, and Lasse H. Pedersen, 2008, Carry trades and currency crashes, *NBER Macroeconomics Annual* (forthcoming).

Brunnermeier, Markus K., and Lasse H. Pedersen, 2008, Market liquidity and funding liquidity, *Review of Financial Studies* (forthcoming).

Caballero, Ricardo J., and Arvind Krishnamurthy, 2003, Excessive dollar debt: Financial development and underinsurance, *Journal of Finance* 58, 867–893.

Cooper, Michael J., Huseyin Gulen, and Michael J. Schill, 2008, Asset growth and the cross-section of stock returns, *Journal of Finance* 63, 1609–1651.

Daniel, Kent, and Sheridan Titman, 2006, Market reactions to tangible and intangible information, *Journal of Finance* 61, 1605–1643.

DeLong, J. Bradford, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann, 1990a, Noise trader risk in financial markets, *Journal of Political Economy* 98, 703–738.

DeLong, J. Bradford, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann, 1990b, Positive feedback investment strategies and destabilizing rational speculation, *Journal of Finance* 45, 379–395.

Diamond, Douglas W., 2004, Presidential address: Committing to commit: Short-term debt when enforcement is costly, *Journal of Finance* 59, 1447–1479.

Diamond, Douglas W., and Raghuram G. Rajan, 2001, Liquidity risk, liquidity creation, and financial fragility: A theory of banking, *Journal of Political Economy* 109, 287–327.

Fama, Eugene F., and Kenneth R. French, 1992, The cross-section of expected returns, *Journal of Finance* 47, 427–465.

Fostel, Ana, and John Geanakoplos, 2008, Leverage cycles and the anxious economy, *American Economic Review* 98, 1211–1244.

French, Kenneth R., 2008, Presidential address: The cost of active investing, *Journal of Finance* 63, 1537–1573.

Friedman, Milton, 1953, The case for flexible exchange rates, in Milton Friedman, ed. *Essays in Positive Economics* (University of Chicago Press, Chicago, IL).

Froot, Kenneth A., and Emil Dabora, 1999, How are stock prices influenced by the location of trade? *Journal of Financial Economics* 53, 189–216.

Gromb, Denis, and Dimitri Vayanos, 2002, Equilibrium and welfare in markets with financially constrained arbitrageurs, *Journal of Financial Economics* 66, 361–407.

Grossman, Sanford J., and Joseph E. Stiglitz, 1980, On the impossibility of informationally efficient markets, *American Economic Review* 70, 393–408.

Hart, Oliver D., and David M. Kreps, 1986, Price destabilizing speculation, *Journal of Political Economy* 94, 927–952.

Hau, Harald, 2008, Global versus local asset pricing: Evidence from arbitrage of the MSCI index change, Working paper, INSEAD.

Hong, Harrison, and Jeremy C. Stein, 1999, A unified theory of underreaction, momentum trading and overreaction in asset markets, *Journal of Finance* 54, 2143–2184.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 93–130.

Kashyap, Anil K., Raghuram G. Rajan, and Jeremy C. Stein, 2008, Rethinking capital regulation, in *Maintaining Stability in a Changing Financial System*, Federal Reserve Bank, Kansas City.

Khandani, Amir E., and Andrew W. Lo, 2007, What happened to the quants in August 2007? Evidence from factors and transactions data, Working paper, MIT.

Kyle, Albert S., and Wei Xiong, 2001, Contagion as a wealth effect, *Journal of Finance* 56, 1401–1440.

Lakonishok, Josef, Andrei Shleifer, and Robert W. Vishny, 1994, Contrarian investment, extrapolation and risk, *Journal of Finance* 49, 1541–1578.

Lehman Brothers, 2000, MSCI confirms free float adjustment and 85% coverage, *Equity Derivatives and Quantitative Research*, December 11.

Lehman, Brothers, 2001, D-day for MSCI, *Equity Derivatives Global*, December 3.

Lehman, Brothers, 2002, Recap of the MSCI May 31, 2002 free float rebalance, *Equity Derivatives Global*, June 10.

Morris, Stephen, and Hyun Song Shin, 2004, Liquidity black holes, *Review of Finance* 8, 1–18.

Rock, Kevin, 1986, Why new issues are underpriced, *Journal of Financial Economics* 15, 187–212.

Shleifer, Andrei, and Robert W. Vishny, 1992, Liquidation values and debt capacity: A market equilibrium approach, *Journal of Finance* 47, 1343–1366.

Shleifer, Andrei, and Robert W. Vishny, 1997, The limits of arbitrage, *Journal of Finance* 52, 35–55.

Sloan, Richard G., 1996, Do stock prices reflect information in accruals and cash flows about future earnings? *Accounting Review* 71, 289–315.

Stein, Jeremy C., 1987, Informational externalities and welfare-reducing speculation, *Journal of Political Economy* 95, 1123–1145.

Stein, Jeremy C., 2005, Why are most funds open-end? Competition and the limits of arbitrage, *Quarterly Journal of Economics* 120, 247–272.