Differences of Opinion, Short-Sales Constraints, and Market Crashes

Harrison Hong
Princeton University

Jeremy C. Stein
Harvard University and NBER

We develop a theory of market crashes based on differences of opinion among investors. Because of short-sales constraints, bearish investors do not initially participate in the market and their information is not revealed in prices. However, if other previously bullish investors bail out of the market, the originally bearish group may become the marginal "support buyers," and more will be learned about their signals. Thus accumulated hidden information comes out during market declines. The model explains a variety of stylized facts about crashes and also makes a distinctive new prediction—that returns will be more negatively skewed conditional on high trading volume.

In this article we address the question of why stock markets may be vulnerable to crashes. To get started we need to articulate precisely what we mean by the word "crash." Our definition of a crash encompasses three distinct elements: 1) A crash is an unusually large movement in stock prices that occurs without a correspondingly large public news event; 2) moreover, this large price change is negative; and 3) a crash is a "contagious" marketwide phenomenon—that is, it involves not just an abrupt decline in the price of a single stock, but rather a highly correlated drop in the prices of an entire class of stocks.

Each of these three elements of our definition can be grounded in a set of robust empirical facts. First, with respect to large price movements in the absence of public news, Cutler, Poterba, and Summers (1989) document that many of the biggest postwar movements in the S&P 500 index—most notably the stock-market break of October 1987—have not been accompanied by any particularly dramatic news events. Similarly Roll (1984, 1988) and French and Roll (1986) demonstrate in various ways that it is hard to explain asset price movements with tangible public information.

The second element of our definition is motivated by a striking empirical asymmetry—the fact that big price changes are more likely to be decreases rather than increases. In other words, stock markets melt down, but they don't melt up. This asymmetry can be measured in a couple of ways. One approach is to look directly at historical stock return data; in this vein it can be noted that of the ten biggest one-day movements in the S&P 500 since 1947, nine were declines.[1] More generally, a large literature documents that stock returns exhibit negative skewness, or, equivalently, "asymmetric volatility"—a tendency for volatility to go up with negative returns.[2]

Alternatively, since gauging the probabilities of extreme moves with historical data is inevitably plagued by "peso problems," one can look to options prices for more information on return distributions. Consider, for example, the pricing of three-month S&P 500 options on January 27, 1999, when Black and Scholes (1973) implied volatility was (i) 39.8% for out-of-the-money puts (strike = 80% of current price); (ii) 27.5% for at-the-money options; and (iii) 17.5% for out-of-the-money calls (strike = 120% of the current price). These prices are obviously at odds with the lognormal distribution assumed in the Black–Scholes model, and can only be rationalized with an implied distribution that is strongly negatively skewed. As shown by Bates (2001), Bakshi, Cao, and Chen (1997), and Dumas, Fleming, and Whaley (1998), this pronounced pattern (often termed a "smirk") in index-option implied volatilities has been the norm since the stock-market crash of October 1987.[3]

The third and final element of our definition of crashes is that they are marketwide phenomena. That is, crashes involve a degree of cross-stock contagion. This notion of contagion corresponds to the empirical observation that the correlation of individual stock returns increases sharply in a falling market [see, e.g., Duffee (1995a)]. Again, the results from historical data are corroborated by options prices. For example, Kelly (1994) writes that "US equity index options exhibit a steep volatility (smirk) while single stock options do not have as steep a (smirk). One explanation . . . is that the market anticipates an increase in correlation during a market correction."

In our effort to develop a theory that can come to grips with all three of these empirical regularities, we focus on the consequences of differences of

[1] Moreover, the one increase—of 9.10% on October 21, 1987—was right on the heels of the 20.47% decline on October 19, and arguably represented a working out of the microstructural distortions created on that chaotic day (jammed phone lines, overwhelmed market makers, unexecuted orders, etc.) rather than an independent, autonomous price change.

[2] Work on skewness and asymmetric volatility includes Pindyck (1984), French, Schwert, and Stambaugh (1987), Nelson (1991), Campbell and Hentschel (1992), Engle and Ng (1993), Glosten, Jagannathan and Runkle (1993), Braun, Nelson and Sunier (1995), Duffee (1995b), Bekaert and Wu (2000), and Wu (2001).

[3] These and other recent articles on options pricing find that one can better fit the index-options data by modeling volatility as a diffusion process that is negatively correlated with the process for stock returns. However, they do not address the question of what economic mechanism might be responsible for the negative correlation.

opinion among investors.[4] We model differences of opinion very simply, by assuming that there are two investors, *A* and *B*, each of whom gets a private signal about a stock's terminal payoff. As a matter of objective reality, each investor's signal contains some useful information. However, *A* only pays attention to his own signal, even if that of *B* is revealed to him in prices, and vice versa. Thus, even without any exogenous noise trading, *A* and *B* will typically have different valuations for the asset.

In addition to investors *A* and *B*, our model also incorporates a class of fully rational, risk-neutral arbitrageurs. These arbitrageurs recognize that the best estimate of the stock's true value is obtained by averaging the signals of *A* and *B*. However, the arbitrageurs may not always get to see both of these signals. This is because we assume—and all our results hinge crucially on this assumption—that investors *A* and *B* face short-sales constraints, and therefore can only take long positions in the stock.

To get a feel for the logic behind our model, imagine that at some time 1, investor *B* gets a pessimistic signal, so that his valuation for the stock at this time lies well below *A*'s. Because of the short-sales constraint, investor *B* will simply sit out of the market, and the only trade will be between investor *A* and the arbitrageurs. The arbitrageurs are rational enough to deduce that *B*'s signal is below *A*'s, but they cannot know exactly by how much. Thus the market price at time 1 impounds *A*'s prior information but does not fully reflect *B*'s time 1 signal.

Next, suppose that at time 2, investor *A* gets a new positive signal. Since *A* continues to be the more optimistic of the two, his new time 2 signal is incorporated into the price, while *B*'s preexisting time 1 signal remains hidden.

Now contrast this with the situation where investor *A* gets a bad signal at time 2. Here things are more complicated, and it is possible that some of *B*'s previously hidden time 1 signal may be revealed at time 2. Intuitively, as *A* bails out of the market at time 2, arbitrageurs will learn something by observing if and at what price *B* steps in and starts being willing to buy. For example, it may be that *B* starts buying after the price drops by only 5% from its time 1 value. In this case, the arbitrageurs learn that *B*'s time 1 signal was not all that bad. But if *B* doesn't step in even after the price drops by 20%, then the arbitrageurs must conclude that *B*'s time 1 signal was more negative than they had previously thought. In other words, the failure of *B* to offer "buying support" in the face of *A*'s selling is additional bad news for the arbitrageurs above and beyond the direct bad news that is inherent in *A*'s desire to sell.

---

[4] Harris and Raviv (1993), Kandel and Pearson (1995), and Odean (1998) are among the recent articles that emphasize the importance of differences of opinion. However, the focus in these articles is primarily on understanding trading volume, not large price movements. See also Harrison and Kreps (1978) and Varian (1989) for related work.

It is easy to see from this discussion how the model captures the first two elements in our definition of a crash. First, note that the price movement at time 2 may be totally out of proportion to the news arrival (i.e., the signal to A) that occurs at this time, since it may also reflect the impact of B's previously hidden signal. In this sense we are quite close in spirit to Romer (1993), who makes the very insightful point that the trading process can cause the endogenous revelation of pent-up private information, and can therefore lead to large price changes based on only small observable contemporaneous news events.[5]

Second—and here we differ sharply from Romer, whose model is inherently symmetric—there is a fundamental asymmetry at work in our framework. When A gets a good signal at time 2, it is revealed in the price, but nothing else is. However, when A gets a bad signal at time 2, not only is this signal revealed, but B's prior hidden information may come out as well. Thus more total information comes out when the market is falling (i.e., when A has a bad signal), which is another way of saying that the biggest observed price movements will be declines.

The one feature of our model that is not readily apparent from the brief discussion above is the one having to do with contagion, or increased correlation among stocks in a downturn. To get at this we have to augment the story so that there are multiple stocks. This opens the possibility that a sell-off in one stock i causes the release of pent-up information that is not only relevant for pricing that stock i, but also for pricing another stock j. Consequently, bad news tends to heighten the correlation among stocks. And of interest is that the price of stock j may now move significantly at a time when there is absolutely no contemporaneous news about its own fundamentals.

In addition to fitting these existing stylized facts, the theory makes further distinctive predictions which allow for "out-of-sample" tests. These predictions have to do with the conditional nature of return asymmetries—that is, the circumstances under which negative skewness in returns will the strongest. When the differences of opinion that set the stage for negative asymmetries are most pronounced, there tends to be abnormally high trading volume. Therefore elevated trading volume should be associated with increased negative skewness, both in the time series and in the cross section.

In empirical work that was initiated after the first draft of this article was completed [Chen, Hong, and Stein (2001)], we develop evidence consistent with these predictions about the conditional nature of skewness. At the same time, however, we also document a fact that, on the face of it, is harder to square with our model: the unconditional *average* skewness of daily returns for individual stocks is positive, in contrast to the significant negative skewness in the returns of the market portfolio. Indeed, one might argue that this

---

[5] See also Caplin and Leahy (1994) for another model in which previously hidden information can be endogenously revealed in large clumps.

fact is particularly at odds with our theory to the extent that individual stocks are harder to short than the market as a whole. As we discuss in detail below, the positive average skewness of individual stocks most likely reflects factors that are left out of the model, such as a tendency for managers to release negative firm-specific information in a gradual piecemeal fashion.

Our theory of crashes can be thought of as "behavioral," in that it relies on less-than-fully rational behavior on the part of investors *A* and *B*. Indeed, the differences of opinion that we model can be interpreted as a form of overconfidence, whereby each investor (incorrectly) thinks his own private signal is more precise than the other's. Or, alternatively, as in Hong and Stein (1999), the differences of opinion can be thought of as reflecting a type of bounded rationality in which investors are simply unable to make inferences from prices. Of course, the usual critique that is applied to these sorts of models is "what happens when one allows for rational arbitrage?" And, in fact, in most models in the behavioral genre, sufficiently risk-tolerant rational arbitrage tends to blunt or even eliminate the impact of the less-rational agents.

In contrast, our results go through even with rational risk-neutral arbitrageurs who can take infinitely long or short positions. This is because the interplay between the arbitrageurs and the less-rational investors is different than in, say, the noise-trader framework of DeLong et al. (1990). In their setting, the less-rational traders have no information about fundamentals, and so the job of the arbitrageurs is just to absorb the additional risk that these noise traders create. In our model, the job of the arbitrageurs is more complicated, because while investors *A* and *B* are not fully rational, they do have access to legitimate private information that the arbitrageurs need. Thus infinite risk tolerance on the part of the rational arbitrageurs is not sufficient to make the model equivalent to one in which everybody behaves fully rationally.[6]

Of course, by making our arbitrageurs risk neutral, we lose the ability to say anything about expected returns—all expected returns in our model are zero, and our implications are only for the higher-order moments of the return distribution. So unlike much of the behavioral finance literature, we do not attempt to speak to the large body of empirical evidence on return predictability. But it is interesting to note that while behavioral models have been used extensively to address the facts on predictability, as well as to explain trading volume, there has been very little serious effort (of which we are aware) to explain market crashes based on behavioral considerations. Ironically, all the best existing models of large price movements are, like Romer (1993), rational models.[7] It is not much of an exaggeration to say

---

[6] The idea that arbitrageurs interact with a class of investors who have valuable information but who overweight this information is also central to Hong and Stein (1999).

[7] We discuss this "rational crash" literature in detail below.

that a state-of-the-art behavioral explanation of a market crash is something along the lines of "there was an abrupt change in investor sentiment."

The remainder of the article is organized as follows. In Section 1 we lay out the assumptions of our model. For simplicity, we consider the case where there is a single traded asset, which can be interpreted either as an individual stock or as the market portfolio. In Section 2 we solve the model and flesh out its implications for the distribution of returns at different horizons. In Section 3 we briefly examine a couple of multiple-asset extensions, which allow us to address issues such as the potential for increased cross-stock correlations in a falling market. In Section 4 we examine the model's empirical content. In Section 5 we discuss the link between our work and previous research on large price movements and/or return asymmetries. Section 6 concludes.

1. The Model

1.1 Timing and information structure

Our model has four dates, which we label times 0, 1, 2, and 3. Initially we consider the case where there is one "stock" that will pay a terminal dividend of D at time 3; it should be stressed that this "stock" can equally well be thought of as the market portfolio. There are three potential traders in the stock: investors A and B, and a group of competitive, risk-neutral rational arbitrageurs. Investors A and B are subject to short-sales constraints, but the arbitrageurs are not.[8] One can interpret the short-sales constraints literally, but they might also be thought of as reflecting institutional restrictions—for example, A and B might be mutual fund managers who, by virtue of their charters or regulation, are deterred from taking short positions.[9]

Investors A and B take turns getting informative signals about the terminal dividend. In particular, at time 1, investor B observes S_B, and next, at time 2, investor A observes S_A. From an objective rational perspective (that of the arbitrageurs), each of these signals is equally informative, as the terminal dividend is given by

$$D = (S_A + S_B)/2 + \epsilon, \tag{1}$$

---

[8] Of importance is that the model does not rest on the assumption that all or even most players are subject to the short-sales constraints. Indeed, the unconstrained risk-neutral arbitrageurs can be seen as representing the vast majority of buying power in the market. All that we really require is that some investors who have significant information be constrained.

[9] A relevant fact in this regard comes from Almazan et al. (2001). They document that roughly 70% of mutual funds explicitly state (in Form N-SAR that they file with the SEC) that they are not permitted to sell short. This is obviously a lower bound on the fraction of funds that never take short positions. Moreover, Koski and Pontiff (1999) find in a study of 679 equity mutual funds that more than 79% of the funds make no use whatsoever of derivatives (either futures or options). Given that derivatives are likely to be the most efficient means for implementing a short position, it would not appear that our approach is founded on an empirically unrealistic premise.

where ϵ is a normally distributed shock with mean zero and variance normalized to one.

As discussed in the introduction, investors A and B each incorrectly believe that only their own signals are informative. This behavioral bias, which can be thought of as a form of overconfidence, induces a difference of opinion among the various agents in the model as to the value of the stock. So, for example, when investor A observes S_A at time 2, he believes that the terminal dividend has an expected value of S_A, irrespective of anything he might be able to infer about S_B.[10] Assuming for simplicity that investor A has CARA utility with a risk aversion coefficient of one, if he is offered the stock at time 2 at a price of p_2, his demand will, in light of the short-sales constraint, be given by

$$Q_A(p_2) = \max[S_A - p_2, \, 0]. \tag{2}$$

Here we are using the lowercase notation p_2 to indicate that we are talking about a "trial" price that may be off the equilibrium path.[11] We will reserve the uppercase notation P_1 and P_2 to refer to the equilibrium prices that are realized at time 1 and time 2, respectively; the significance of this distinction will become apparent shortly. Similarly investor B's demand for the stock at time t $(t = 1, 2)$, if he is offered the stock at a price of p_t, will be[12]

$$Q_B(p_t) = \max[S_B - p_t, \, 0]. \tag{3}$$

Prior to being realized at time 1, S_B is uniformly distributed on the interval $[0, 2V]$. Thus the rational expectation of S_B as of time 0 is $E_0[S_B] = V$. Prior to being realized at time 2, S_A is uniformly and independently distributed on $[H, 2V + H]$, so that $E_0[S_A] = E_1[S_A] = V + H$. Note that V can be interpreted as a measure of the variance of the news that is received by the investors, while H can be thought of as an ex ante measure of the heterogeneity of their opinions. In what follows, we assume that $0 \leq H \leq 2V$. This implies that investor B—who moves first—is on average more bearish than investor A. This assumption is not crucial to our results. Indeed, we discuss in Section 2.4.1 below how the results generalize when we reformulate the model so that the bearish investor B moves first with probability one-half, while the bullish investor A also moves first with probability one-half. The reason that we begin by restricting ourselves to the case where B always

---

[10] The degree of overconfidence need not be as extreme as we are assuming here. All that we really need is that A does not adjust his valuation all the way to the rational level given in Equation (1) upon learning B's signal, and vice versa.

[11] We are treating A and B as price takers. So it may be more accurate to think of each of them as corresponding to a group of competitive investors who all get the same signal.

[12] For simplicity, we are assuming that investor B's demand for the stock at time 1 depends only on his expectation of the terminal dividend, and not on the price that he expects to prevail at time 2. This assumption is not at all critical for our results.

moves first is that, as will become clear shortly, this case allows us to high-light the central intuition of the model, while greatly reducing the complexity of the analysis.

1.2 The price-setting mechanism

We now turn to the determination of prices at the various dates. Note that because of the risk neutrality of the arbitrageurs, we can without loss of generality set the supply of the stock to zero. It is also easy to see that the price at time 0 is given by

$$P_0 = V + H/2. \tag{4}$$

This is just the arbitrageurs' ex ante expectation of the terminal dividend before either A or B have received their signals. But once these signals begin to be realized, at times 1 and 2, the issue of price setting becomes a bit more complicated, and we have to be clear about the mechanism that is used.

We assume the following setup at times 1 and 2. Investors A and B, along with the arbitrageurs, are all together in a room with an auctioneer. Any time the auctioneer announces a trial price p_t, the participants respond by calling out their demands. Because of the short-sales constraints, investors A and B only call out something if their demands are positive; otherwise they are silent. The arbitrageurs, who face no short-sales constraint, are free to call out either positive or negative demands. It is important that the arbitrageurs are able to observe any demands called out by investors A and B.

The auctioneer follows a simple mechanical rule, which could be carried out by a computer. He starts by announcing a "high" trial price, say $2V + H$, which is known to be higher than anybody's highest possible valuation. At this high price, the net excess demand for the stock is certain to be negative. The auctioneer then gradually begins to adjust the price. His adjustment rule is that as long as the excess demand remains negative, he lowers the price. Conversely, if he ever reaches a point where the excess demand is positive, he raises the price. This process continues until the market clears—that is, until the auctioneer finds a price such that the excess demand for the stock is zero.

As will become clear below, this auction mechanism provides us with a simple way to determine a unique equilibrium price at each date. Moreover, the equilbrium will have the intuitive property that whichever investor (A or B) has the more positive signal at a given date will be long the stock, and his signal will be fully revealed. In contrast, the investor with the less positive signal will not own any shares in equilibrium, and his signal may or may not be revealed in the course of the auction. Although it is obviously some-thing of a modeling contrivance, we do not think that our auction scheme is too unrealistic. In fact, it resembles quite closely the opening procedures

used in several major stock markets, including the Paris Bourse, the Toronto Exchange, and the New York Stock Exchange.[13]

1.3 The rational expectations benchmark

Before solving the model with differences of opinion, we digress briefly and consider the benchmark case where all the players in the model are fully rational. In this benchmark case, investors A and B, like the arbitrageurs, recognize that the best estimate of the terminal dividend (conditional on knowing S_A and S_B) is given by $(S_A + S_B)/2$ rather than just by their own private signals. This results in the following outcome:

Proposition 1. *When investors A and B are fully rational, the short-sales constraint does not bind. Prices fully reflect all information as soon as it becomes available to investors*:

$$P_1 = (V + H + S_B)/2; \tag{5a}$$

$$P_2 = (S_A + S_B)/2. \tag{5b}$$

Consequently, returns are symmetrically distributed at time 1 *and time* 2. *Returns are also homoscedastic—that is, they have the same variance at time* 1 *as at time* 2.

To see the logic of the proof, consider time 1, and suppose that investor B's information has not yet come out during the auction process. This implies that investor B's estimate of the terminal dividend is lower than any trial price p_1 that has been announced. At the same time, any market-clearing price P_1 must equal the risk-neutral arbitrageurs' estimate of the terminal dividend. And the arbitrageurs recognize that investor B is rational and strictly better informed than they are at time 1. Thus as long as S_B has not been revealed, the arbitrageurs know that the trial price p_1 is too high, and the market cannot clear. Similar reasoning establishes that the market cannot clear at time 2 unless S_A has been revealed.

Proposition 1 is significant because it highlights the key role that differences of opinion play in our model. The results on return asymmetries and heteroscedasticity that we obtain below are not driven solely by the short-sales constraint; rather the short-sales constraint must interact with the differences of opinion to generate anything interesting.[14]

---

[13] The Paris Bourse would seem to be especially close to what we have in mind: during a preopening period, trial clearing prices are transmitted to some traders along with information on excess demand at those prices, and the traders can revise their orders multiple times before a market clearing price is established. See Domowitz and Madhavan (1998) for details.

[14] This feature distinguishes our model from that of Diamond and Verrecchia (1987), where short-sales constraints matter even in a setting where everybody is rational. Loosely speaking, the difference arises because our price-setting mechanism allows for more information sharing among traders at a given point in time than theirs, which in turn gives the rationality assumption more bite.

2. Solving the Model with Differences of Opinion

2.1 Time 1: the potential for hidden information

We now turn back to the situation where there are differences of opinion. Now it is possible that an investor's signal may not be revealed in equilibrium, if he is sufficiently pessimistic. Let us first examine what happens at time 1, when the only private information is held by investor B. We can distinguish two possible cases:

Case 1. *Investor B's information is revealed, in which case*

$$P_1 = (V + H + S_B)/2. \tag{6}$$

Case 2. *Investor B's information remains hidden, in which case*

$$P_1 = (V + H)/2 + E_1[S_B|\text{NR}]/2, \tag{7}$$

where $E_1[S_B|\text{NR}]$ is the time 1 conditional expectation of S_B, given that S_B has not been revealed.

When can Case 2 occur? Given our auction mechanism, a necessary condition for S_B to remain hidden is that $S_B \leq P_1$. In words, investor B's valuation must not exceed the market-clearing price, or otherwise he would have called out a nonzero demand during the auction, thereby tipping his signal to the arbitrageurs. Holding the arbitrageurs' conjectures fixed, the necessary condition becomes harder to satisfy the higher is S_B. This suggests that there will be a cutoff value of S_B—which we denote by S_B^*—such that if S_B lies above S_B^*, the equilibrium must involve revelation of S_B.

It is easy to establish what the value of S_B^* must be. If there is revelation for all values of $S_B > S_B^*$, then the expected value of S_B conditional on no revelation, $E_1[S_B|\text{NR}]$, must equal $S_B^*/2$. This implies that the price in Case 2 is given by

$$P_1 = (V + H)/2 + S_B^*/4. \tag{8}$$

But we can only be in Case 2 if $S_B \leq P_1$. So one solves for the cutoff S_B^* by setting it equal to P_1, which gives us:

Lemma 1. *Let the cutoff value for S_B be*

$$S_B^* = 2(V + H)/3. \tag{9}$$

Then for all values of $S_B > S_B^*$, there must be revelation of S_B—that is, we must be in Case 1.

It is worthwhile to map out specifically how the auction mechanism works in Case 1, when $S_B > S_B^*$, and hence S_B is revealed. There are two qualitatively distinct scenarios. In the first, S_B is "very high;" in particular

$S_B > (V + H)$, which means that investor B's valuation is higher even than the ex ante expectation of A's valuation. (Note that this can only occur if we make the assumption that $H < V$.) As the auctioneer starts to work the trial price down, initially all he hears are sell orders from the arbitrageurs, as the trial price is above everybody's valuation. When the auctioneer gets to a trial price $p_1 = S_B$, investor B calls out, revealing his signal. At this point, the arbitrageurs become fully informed. They recognize that the true value of the stock is $(V + H + S_B)/2$, so at $p_1 = S_B$, they continue to want to sell. Thus the price keeps dropping until it hits $(V + H + S_B)/2$, at which point the market clears. Observe that in this scenario, the bullish investor B is long the stock in equilibrium.

In the second scenario of Case 1, S_B is only "moderately high;" that is, $S_B^* < S_B \leq (V + H)$. Now the auctioneer's trial price can drop further with investor B staying silent. For any trial price p_1 in this silent region, and below $2V$, the arbitrageurs' conditional estimate of the terminal payoff is just $(V + H)/2 + E[S_B | S_B \leq p_1]/2 = (V + H)/2 + p_1/4$. This implies that the longer investor B stays quiet in the face of dropping trial prices, the lower the arbitrageurs' estimate drops. Moreover, as long as this estimate remains below the trial price of p_1, the risk-neutral arbitrageurs have infinite negative demand. This in turn causes the auctioneer to move p_1 down further. When p_1 hits S_B, investor B calls out, thereby revealing his signal. This is good news for the arbitrageurs—they learn that they are in Case 1 rather than Case 2—so their estimate of the value jumps discretely, to $(V + H + S_B)/2$. As a result, there is now positive excess demand, and the auctioneer has to raise the price back up to meet the arbitrageurs' new estimate, at which point the market finally clears.

A subtle point about this second scenario is that S_B is revealed through the auction process even though, in equilibrium, investor B ends up holding no shares. This is because in this scenario, the trial price at some point necessarily falls below the ultimate equilibrium price, causing investor B to call out a demand and reveal his signal.

If $S_B \leq S_B^*$, however, S_B can remain concealed. More precisely, given our auction mechanism, we can show:

Lemma 2. *For all values of $S_B \leq S_B^*$, the unique equilibrium involves the "pooling" outcome of Case 2, where S_B remains hidden, and where*

$$P_1 = (V + H)/2 + S_B^*/4 = 2(V + H)/3. \tag{10}$$

The ex ante probability of winding up in this pooling equilibrium is $(V + H)/3V$. That is, pooling is more likely when there is more ex ante heterogeneity in opinions, as measured by the parameter H.

In this case, the auction proceeds as follows. The auctioneer works the trial price p_1 down, as before. But this time, before p_1 falls to S_B, and hence

before investor B calls out, the trial price hits S_B^*. At $p_1 = S_B^*$, the arbitrageurs' estimate of value, $(V + H)/2 + p_1/4$, equals the prevailing trial price. So the market clears before investor B ever gets in.

It should be pointed out that the uniqueness of the pooling equilibrium for $S_B \leq S_B^*$ is a consequence of our assumptions about the adjustment rule followed by the auctioneer. To see why, suppose the auctioneer was not restricted to adjusting prices gradually, but instead could discontinuously announce a trial price of zero. At this point, investor B would always call out a demand, for any value of S_B. That is, S_B would always be revealed, no matter how low. Thus our auction mechanism, while it is arguably reasonable, is also critical to establishing the central feature of our model—that some information may remain hidden at time 1.[15]

2.2 Time 2: previously hidden information may be revealed

The bottom line from our analysis of time 1 is that if S_B is low enough, it may not be immediately revealed. Now we move to time 2. The primary goal here is to show that a low draw of investor A's signal, S_A, may cause further information on S_B to come out. So naturally, much of our focus will be on that branch of the time 1 tree where we were in Case 2, and S_B was hidden. However, because we ultimately want to be able to provide a complete description of the distribution of returns at both time 1 and time 2, we also need to fill in what happens along the less interesting branch of the tree where there was no hidden information at time 1—that is, where we were previously in Case 1. This is where we begin.

2.2.1 Case 1: *B*'s signal was revealed at time 1.
If S_B has already been revealed, the analysis at time 2 is very similar to that at time 1. If the signal of investor A, S_A, is relatively high, it will also be revealed at time 2. However, if it is sufficiently low, it may remain hidden. In particular we can show

Lemma 3. *Assume that S_B has been revealed at time* 1. *Let the cutoff value for S_A be*

$$S_A^* = (2S_B + H)/3. \tag{11}$$

For all values of $S_A > S_A^*$, S_A is also revealed at time 2, *and*

$$P_2 = (S_A + S_B)/2. \tag{12}$$

*We call this **Case** 1.A.*

Lemma 4. *Assume that S_B has been revealed at time 1. If $S_A \leq S_A^*$, equilibrium at time 2 involves a pooling outcome where S_A remains hidden, and where*

$$P_2 = S_B/2 + (H + S_A^*)/4 = (2S_B + H)/3. \tag{13}$$

*We call this **Case 1**.**B**.*

2.2.2 Case 2: B's signal was hidden at time 1. If S_B has not yet been revealed, the analysis at time 2 is more interesting; this is where the heart of our model lies. The results in this case can be characterized by four lemmas. These four lemmas, which are proven in the appendix, collectively provide a complete characterization of the possible time 2 outcomes for all parameter values.

First, if investor A gets good news at time 2—that is, if S_A turns out to be above its ex ante expectation—no further information on S_B comes out. That is,

Lemma 5. *Assume that S_B was hidden at time 1. If $S_A \geq (V + H)$, then S_A is revealed, and S_B continues to pool below the old time 1 cutoff of S_B^*. The price in this case is given by*

$$P_2 = S_A/2 + S_B^*/4 = S_A/2 + (V + H)/6. \tag{14}$$

*We call this **Case 2**.**A**.*

This makes intuitive sense; given that investor B was too pessimistic relative to the prior on S_A to get into the market and tip his signal at time 1, he certainly won't get in at time 2 if A becomes even more optimistic and the gap between A's and B's valuations widens.

Second, if investor A gets a bad signal at time 2, some further information on S_B will come out. Of importance, however, is that this need not imply that S_B is fully revealed. Instead, it is possible that S_B will still pool, but inside a lower portion of its support. That is, it may be learned that $S_B \leq S_B^{**}$, where S_B^{**} is a new cutoff level that is below S_B^*. This clearly represents a sharpening of the market's information on S_B, but it is not total revelation. Moreover, in this setting, one can meaningfully talk about "how much more" information on S_B has come out—the lower the new cutoff S_B^{**}, the more has been learned. More precisely, we have

Lemma 6. *Assume that S_B was hidden at time 1, and also that $S_A < (V + H)$. Let the new cutoff*

$$S_B^{**} = 2S_A/3. \tag{15}$$

If $S_B \leq S_B^{**}$, then S_A is revealed, and S_B pools below the new cutoff of S_B^{**}. The price in this case is given by

$$P_2 = S_A/2 + S_B^{**}/4 = 2S_A/3. \qquad (16)$$

We call this **Case 2.B**.

Thus, if S_B is small enough relative to S_A, it may still remain partially hidden at time 2. On the other hand, if S_B exceeds the new cutoff S_B^{**}, it will be fully revealed. In fact, there are two distinct scenarios in which S_B is fully revealed. First, S_B may be fully revealed, while S_A is hidden below a cutoff value of S_A^*:

Lemma 7. *Assume that S_B was hidden at time* 1, *and also that $S_A < (V + H)$. As in Lemma* 3, *let the cutoff on S_A be*

$$S_A^* = (2S_B + H)/3. \qquad (17)$$

If $S_A \leq S_A^*$, and simultaneously $S_B > H$, then S_A pools below S_A^*, while S_B is fully revealed. The price in this case is given by

$$P_2 = S_B/2 + (H + S_A^*)/4 = (2S_B + H)/3. \qquad (18)$$

We call this **Case 2.C**

Alternatively, both S_B and S_A can be revealed.

Lemma 8. *Assume that S_B was hidden at time* 1. *For any parameter values not already covered in Lemmas* 5–7, *both S_A and S_B are fully revealed at time* 2. *The price in this case is given by*

$$P_2 = (S_A + S_B)/2. \qquad (19)$$

We call this **Case 2.D**

The key point that emerges from the analysis is that, holding fixed the actual realization of S_B, more information on S_B comes out the lower is S_A. This shows up in two ways. First, for a lower S_A, S_B is more likely to be fully revealed. Second, even if it is not fully revealed, a lower value of S_A implies that S_B will remain hidden in a smaller portion of the lower support of its distribution.

We think that this feature of the model—the time 2 link between the realization of S_A and the amount of new information that comes out on S_B— best embodies the central economic intuition that we are trying to capture. Essentially our story is one in which a change of heart on the part of a previously optimistic investor (A) tests the resolve of another previously sidelined investor (B). The extent to which B is willing to step in and offer buying support as A bails out of the market is important data to the arbitrageurs.

And the more completely *A* bails, the more informative is the experiment conducted on the ostensible support buyers.

Figure 1 provides a compact illustration of all of our results to this point. It shows how the entire parameter space can be partitioned into six regions, corresponding to our Cases 1.A, 1.B, 2.A, 2.B, 2.C, and 2.D. Having done this partitioning—and knowing the equilibrium prices P_1 and P_2 that arise in each region—we can now make a variety of statements about the distributional properties of returns. For example, as we will demonstrate shortly, it is a straightforward task to compute the skewness of returns at various horizons, simply by taking the appropriate integrals over the different regions.

2.3 Implications for return asymmetries at different horizons
2.3.1 Asymmetries in "big moves."   One simple, nonparametric way of thinking about asymmetries in the distribution of extreme returns is to calculate
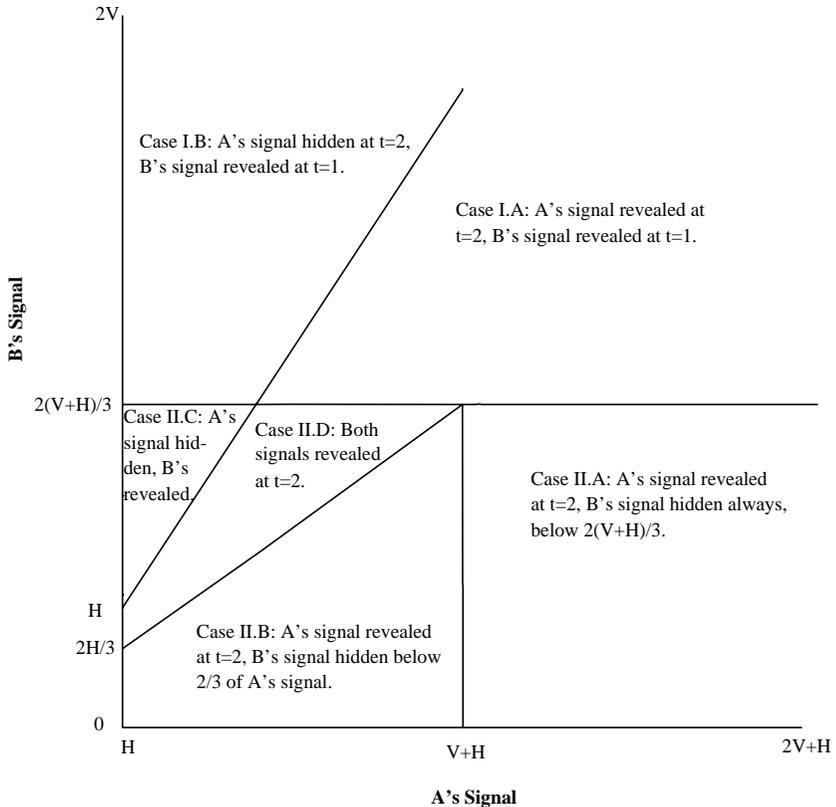

Figure 1
Partition of equilibrium outcomes depending on *A*'s and *B*'s signals, S_A and S_B, respectively
Graphical classification of Cases 1.A, 1.B, 2.A, 2.B, 2.C, and 2.D characterized in Lemmas 3–8.

the largest possible moves—either up or down—that can occur at times 1 and 2. Consider time 1 first, and define the time 1 return

$$R_1 = P_1 - P_0. \tag{20}$$

Given our earlier results, it is easy to show that the biggest possible up-move at time 1, which we denote by $BIG^U(R_1)$, equals $V/2$. This occurs when S_B attains its highest value, and is fully revealed at time 1. In contrast, the biggest possible down-move (in absolute value terms), $BIG^D(R_1)$, is given by $V/3 - H/6$, which is strictly smaller than $BIG^U(R_1)$. Down-moves are less extreme at time 1 because if S_B is very low, it is not fully revealed. Thus the asymmetry at time 1 is the opposite of what we are looking for—it suggests that the largest price movements will be increases, not decreases. Again, this is a direct consequence of the fact that bad news is hidden at time 1.

Next, consider returns at time 2,

$$R_2 = P_2 - P_1. \tag{21}$$

As before, the biggest possible up-move, now denoted $BIG^U(R_2)$, equals $V/2$. This reflects the highest possible realization of S_A, and the observation that when S_A embodies good news, no further information on S_B comes out. But now the biggest possible down-move, $BIG^D(R_2)$, is given by $2V/3$, which is strictly larger than $BIG^U(R_2)$. The reason the down-move can be more extreme is that it represents not only the full revelation of the lowest possible S_A, but also further news about S_B, a piece of information which had been hidden from the market before time 2.

Thus while returns at time 1 are suggestive of a positive asymmetry in the distribution, returns at time 2 are suggestive of a negative asymmetry. Moreover, the effect at time 2 is in a sense stronger, because the variance of returns at this time is greater. One way to express this is as follows:

Proposition 2. *Taking into account both R_1 and R_2, the overall largest possible one-period return occurs on a down-move.*

So in an unconditional sense, it is indeed accurate to say that the distribution of extreme returns is characterized by a negative asymmetry—the biggest movements in the stock price will be decreases. This property of the model corresponds closely to the historical facts discussed in the introduction.

2.3.2 Skewness. An alternative way to measure asymmetries in the return distribution is to calculate the skewness, or third moment, of the distribution. As mentioned earlier, these skewness calculations are conceptually straightforward, though they involve fairly laborious integration. With the help of the computer program Mathematica, we are able to solve everything in closed form, and the results that we report below are based on the properties of these closed-form solutions. All details are in the appendix.

Again, we begin by considering the properties of the time 1 return, R_1. Analogous to our result with big moves, this return is positively skewed:

Lemma 9. *For all values of H and V, $E[R_1^3] \geq 0$.*

Moving to the time 2 return, R_2, we find that things are a little more subtle. Conditional on S_B having been hidden at time 1, R_2 is indeed negatively skewed; this is where the intuiton from our big moves analysis carries over directly. But it turns out that conditional on S_B having been revealed at time 1, R_2 is actually slightly positively skewed, for exactly the same reasons that R_1 is positively skewed—there may be some hiding of bad news, in this case bad news about S_A. (This positive skewness effect at time 2 comes from Case 1.B.) Putting it all together, it turns out that, from an unconditional perspective, R_2 will be negatively skewed for all but the smallest values of H. More precisely, we have

Lemma 10. *Conditional on being in Case 2, $E[R_2^3|\text{Case 2}] < 0$. Conditional on being in Case 1, $E[R_2^3|\text{Case 1}] > 0$. Unconditionally, $E[R_2^3]$ is monotonically decreasing in the ratio H/V, and is negative for values of $H/V > .38$.*

Of course, from an empirical perspective, it is more helpful to be able to make statements that do not depend on whether we take the perspective of time 1 or time 2.[16] In this spirit, we have

Proposition 3. *Define the overall unconditional skewness of short-horizon returns to be*

$$E[R_s^3] = \frac{E[R_1^3] + E[R_2^3]}{2}. \tag{22}$$

Our model has the property that $E[R_s^3] < 0$ for values of $H/V > 1.69$.

Thus one of our central results is that short-horizon returns will, in an unconditional sense, be negatively skewed as long as there is enough ex ante heterogeneity in investors' opinions—that is, as long as H is large enough relative to V. Figure 2 illustrates the results of Lemmas 9 and 10, along with those of Proposition 3, showing how our various conditional and unconditional measures of short-horizon skewness vary with the ratio H/V.

To see intuitively why a high value of H/V necessarily leads to negative skewness, consider the limiting situation where $H = 2V$. As we saw earlier (Lemma 2), in this situation Case 1 disappears, and we are always in Case 2,

---

[16] We are ignoring R_3, the return from time 2 to time 3, in our skewness calculations. Given our assumption of uniform distributions, this is of no consequence—R_3 will always be symmetrically distributed and hence contribute nothing to overall skewness. This is because any remaining information that comes out at time 3 is just a draw from a truncated uniform distribution, which is itself still uniform, and hence symmetric. This feature no longer holds with alternative distributions, as we discuss in Section 2.4.3 below.
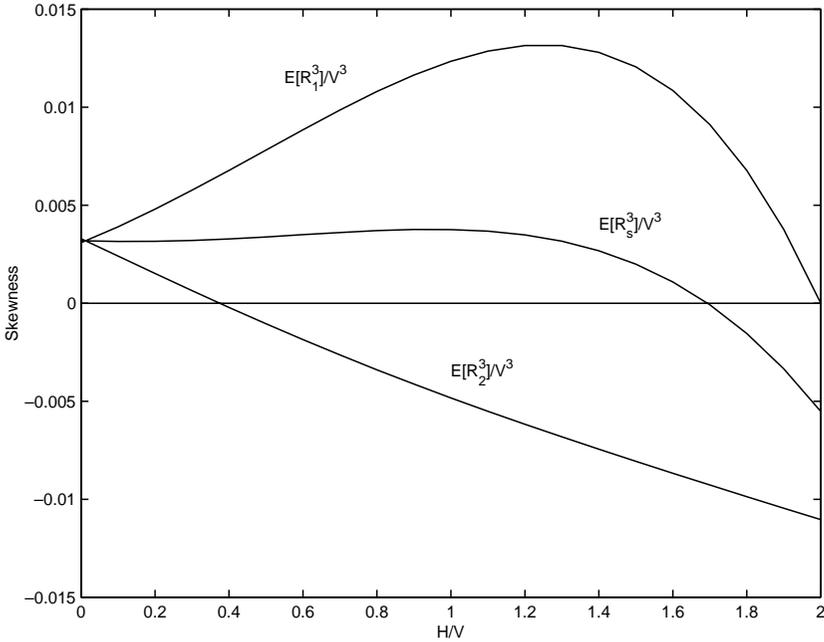
Figure 2
Skewness and differences of opinion
Plot of various skewness measures against a measure of differences of opinion, H/V.

where S_B is hidden at time 1. Consequently prices do not move at all at time 1, so the positive skewness in time 1 returns from Lemma 9 drops out of the picture. All we are left with is the negative Case 2 skewness at time 2 from Lemma 10.

It is useful to pause and ask why these results for skewness appear to be less decisive than those for big moves. Recall that with regard to big moves, we have the sharp conclusion that the largest possible move is always a decline, irrespective of the value of H. In contrast, with skewness, it seems that we need to put some restrictions on H/V to get a clear-cut negative asymmetry.

This divergence reflects the fact that our model embodies two competing effects: a hiding-of-bad-news effect at time 1 that gives rise to a positive asymmetry, and a revelation-of-news effect at time 2 that generates a negative asymmetry. The latter effect always dominates when the metric is big moves, but not necessarily when the metric is skewness, since skewness is influenced in part by returns that are not as far out in the tails of the distribution. To put it another way: we could in principle calculate higher-order odd moments of the return distribution—for example, the fifth moment, the seventh moment, etc. These higher-order moments would be more heavily influenced by the

action far out in the tails, and we conjecture that they would be more likely to be unconditionally negative at short horizons, for a wider range of values of H/V. Nevertheless, the concept of skewness is still an attractive one to focus on, since it is intuitive, easy to calculate (in our model), and allows us to map our findings into the large body of existing evidence that is based on this parametric measure.

In this spirit, another empirically relevant thought experiment is to ask how skewness varies with the return horizon. We begin by defining a scaled measure of medium-horizon returns:

$$R_m = (R_1 + R_2)/\sqrt{2}. \tag{23}$$

The measure is scaled so that, in the rational expectations benchmark, medium-horizon returns have the same variance as short-horizon returns; this sort of adjustment is necessary if we are to make meaningful comparisons of skewness across horizons.[17] With the definition in hand, we can establish the following:

Proposition 4. *For values of $H/V > 1.84$, $E[R_m^3] > E[R_s^3]$. That is, medium-horizon skewness is less negative than short-horizon skewness when there are sufficiently large differences of opinion.*

This result is driven by the following simple logic. As we lengthen the horizon over which returns are calculated, the potential for prices to move very sharply downward in a short interval (between time 1 and time 2) carries less weight, and therefore contributes less to negative skewness. Proposition 4 also squares nicely with the available evidence. For example, Bakshi, Cao, and Chen (1997) and Derman (1999) find that the magnitude of the "smirk" in S&P 500 index option implied volatilities—the extent to which implied volatilities for out-of-the-money puts exceed those for out-of-the-money calls—is a decreasing function of the maturity of the options. Thus the options market is suggesting that negative return skewness in the S&P 500 index diminishes with the horizon over which returns are measured.

2.3.3 Trading volume and conditional skewness. Although we have not emphasized it to this point, our model—like any model incorporating differences of opinion—has straightforward implications for trading volume.[18] Simply put, when the heterogeneity parameter H is larger, there will tend to be more turnover. But since the heterogeneity parameter also governs the degree of skewness, we have a novel prediction about conditional skewness: that higher trading volume is associated with more negative skewness.

---

[17] Without the scaling, there would be a strong tendency for longer-horizon returns, due to their greater variance, to have higher raw third moments.

[18] See, for example, Harris and Raviv (1993), Kandel and Pearson (1995), and Odean (1998) for other models where differences of opinion drive trading volume.

To show more clearly why this is so, we consider a situation where $H > V$.[19] We continue to assume that, as of time 0, investors A and B have no initial endowment of the stock. Now let us ask what cumulative trading volume is at times 1 and 2.[20]

First, observe that the only potential buyer at time 1 is B, since A has not yet received his signal. But the assumption that $H > V$ implies that $S_B < P_1$, so B never buys in equilibrium at time 1. Hence the only trade is at time 2. At this time, if A has the higher valuation, trading volume will be proportional to $(S_A - P_2)$. Conversely, if B has the higher valuation, trading volume will be proportional to $(S_B - P_2)$. This logic can be used to show that overall expected trading volume will be proportional to the expected value of $|S_A - S_B|/2$ (see the proof of Lemma 11 in the appendix). This is very intuitive—trading volume arises to the extent that A's and B's signals differ from one another.

In the appendix, we use this observation to prove

Lemma 11. *For $H > V$, all trades take place at time 2, and expected trading volume is given by $kE[|S_A - S_B|/2] = kH/4$, where k is a positive constant.*

The implications for conditional skewness follow immediately from the lemma, since, as can be seen in Figure 2, $E[R_s^3]$ is monotonically decreasing in H for $H > V$. Hence we have established

Proposition 5. *For $H > V$, the degree of negative skewness in short-horizon returns is increasing in trading volume.*

Proposition 5 can in principle be tested quite directly, using either time-series data on the aggregate market or data on individual stocks. Moreover, it would appear to be particularly useful in constructing a sharp, out-of-sample test of our theory. For while there are other models that can deliver negative asymmetries in returns, we are not aware of any that link these negative asymmetries to trading volume. We return to these issues in more detail below.

2.4 Robustness issues
2.4.1 Investors A and B have equal probability of moving first.
The analysis thus far has assumed that the more bearish investor B moves first, at time 1, while the more bullish investor A does not move until time 2. This assumption is a natural way to highlight the intuition of our model, since the model relies on information being hidden at time 1, and B's information

---

[19] This parametric restriction simplifies the analysis but is not strictly necessary for the results that follow.

[20] Implicitly, we are assuming that A and B are not active in the market at time 0 and that prices at this time are set (without any trading volume) by the arbitrageurs, who all agree on what the stock is worth ex ante.

is more likely to be hidden, given that his valuations are on average lower than A's.

Nevertheless, it is important to make it clear that this asymmetric ordering is not responsible for our main results. We have redone everything with an alternative formulation of the model in which B moves first with probability one-half and A moves first with probability one-half. Although we do not present this analysis in detail here because it is very lengthy—it effectively involves reworking every one of the preceding lemmas and propositions for the case where A moves first—the bottom line is that none of our key conclusions is materially changed. (Details are available from the authors upon request.) For example, Proposition 2 about big moves continues to go through exactly as stated. And Proposition 3 is actually strengthened slightly: the conclusion is now that $E[R_s^3] < 0$ for values of $H/V > 1.30$, whereas before, we required $H/V > 1.69$ to generate negative skewness.

2.4.2 Alternative auction mechanisms. One concern about our model is that the results seem to depend crucially on the form of the auction mechanism, which allows low signals to remain hidden. Although this particular mechanism is attractive as a modeling device because of its simplicity and tractability, there are alternative ways to generate the key feature of hidden information at time 1.

For example, suppose that it is costly for investors to show up at the auction in the first place. In this case, they will not want to show up if there is no chance that they will wind up buying any stock. Now consider the decision facing a B investor at time 1 who has drawn a relatively low signal S_B. Given his priors on the distribution of S_A, and his understanding of the model, he may be able to anticipate that any market-clearing price P_1 would certainly lie above S_B, regardless of what auction mechanism is used, and regardless of how much of S_B is revealed in the process. Thus it will not be worth it for him to attend the auction, and S_B cannot possibly be revealed in equilibrium, no matter how the auction is conducted.

2.4.3 Distributions other than the uniform. Clearly the assumption that signals are uniformly distributed greatly simplifies the analysis. But to what extent does it color our conclusions? Although we have not been able to fully redo the model with alternative distributions, we offer the following conjectures.

First, the nonparametric results for big moves summarized in Proposition 2 would seem to be robust and insensitive to the specific signal distribution that is assumed. These results only rely on the most basic intuition from the model—that more previously hidden information is revealed when bad news comes out at time 2. As we have shown, these results already hold without regard to the value of H.

Second, the precise parametric results for skewness seen in, for example, Proposition 3, are no doubt more likely to be influenced by distributional assumptions—a clue to this is the sensitivity of skewness to the parameter H. But our intuition suggests that many alternative distributions may actually lead to more pronounced negative skewness. To see why, note that with a uniform signal distribution, the only way to generate negative skewness is through our endogenous information revelation effect—previously hidden information has to be forced out by the arrival of other bad news. In contrast, if previously hidden information simply comes out through, say, an earnings announcement, there will be no skewness, because a truncated uniform is still uniform and hence symmetric.[21]

If signals were instead normally distributed, it seems that there would be more potential for negative skewness. In particular, consider what would happen if S_B were hidden at time 1, and did not come out at time 2, because there was a high draw of S_A. Now returns at time 3 would be negatively skewed because the conditional distribution of S_B—given that it is a normal truncated from above—is left-skewed. Although obviously far from a complete treatment, this discussion suggests that, by assuming uniformly distributed signals, our model is actually conservative, in the sense that it has a harder time generating negative skewness than it otherwise might.

3. Multiple Stocks and Contagion

As discussed in the introduction, there is evidence that individual stocks become more highly correlated with one another during market declines. This effect might be seen as indicative of a form of "contagion." We now discuss two variations on our model which illustrate how such an effect might arise.

3.1 Differences of opinion about the market factor

One very easy way to generate increased cross-stock correlations in a market downturn is simply to argue that the returns on the market portfolio itself are, for some reason, negatively skewed. To be more precise, suppose that there are a large number of stocks and that the return on any given stock i obeys a one-factor structure:

$$R_i = R_M + Z_i, \tag{24}$$

where R_i is the return on stock i, R_M is the mean-zero return on the market factor, and Z_i is a mean-zero idiosyncratic component that is independently and identically distributed across all stocks and independent of R_M.

---

[21] This is why the distribution of the third-period return R_3 is always symmetric in our model. This feature distinguishes our model from Diamond and Verrecchia (1987), where returns around earnings announcements can be negatively skewed.

For a given realization of R_M, we define $\hat{\sigma}_{ij}$ to be the average value (across all pairs of stocks) of the sample estimator of the covariance between the returns on any two stocks i and j:

$$\hat{\sigma}_{ij} = E[R_i R_j | R_M]. \tag{25}$$

Similarly we define $\hat{\rho}_{ij}$ as the average value of the sample estimator of the correlation between the returns on any two stocks i and j,

$$\hat{\rho}_{ij} = \frac{E[R_i R_j | R_M]}{\sqrt{E[R_i^2 | R_M] E[R_j^2 | R_M]}}. \tag{26}$$

Note that both $\hat{\sigma}_{ij}$ and $\hat{\rho}_{ij}$ are random variables that depend on the realization of R_M.

In the appendix, we show that

Proposition 6. *If the return on the market factor is negatively skewed,* $E[R_M^3] < 0$, *then (i)* $cov(\hat{\sigma}_{ij}, R_M) < 0$, *and (ii)* $cov(\hat{\rho}_{ij}, R_M) < 0$.

The logic is straightforward. Conditional on a large movement in the market factor, individual stock returns are highly correlated—this is true even if the market factor is symmetrically distributed. But if the market factor is negatively skewed, then a large movement in it is more likely to occur on a decline. So on average, declines in the market factor correspond to increased values of $\hat{\sigma}_{ij}$ and $\hat{\rho}_{ij}$.

Therefore, if one assumes that the market factor is itself a traded asset, and that our model can be applied directly to it, the result about increased cross-stock correlations in a downturn will follow (at short horizons) so long as there are sufficient differences of opinion about the market factor that the conditions of Proposition 3 apply. In other words, if for the market factor, we have that $H/V > 1.69$, then the market factor will exhibit negative skewness at short horizons, and cross-stock correlations will covary negatively with market returns.

3.2 Idiosyncratic shocks to stock *i* spill over to stock *j*

In the preceding setup, "contagion" arises through shocks to the market factor. That is, when there is a large drop in the market factor, all stocks tend to fall together. Although this approach is adequate for fitting the empirical facts about increased correlations in downturns, it may not quite capture the economic intuition that many people have about contagion. In this regard, perhaps a sharper definition of contagion is the idea that when there is bad firm-specific news about one stock *i* at some time *t*, this causes a decline

in the price of another stock j, even though the time t news has absolutely nothing to do with stock j's own fundamentals.[22]

To capture this idea, we consider an extension of our model where there are two stocks. Each stock has its own groups of investors, A_i and B_i for stock i, and analogously for stock j. Each stock also has its own group of arbitrageurs. To keep things simple we focus on the polar case where for both stocks $H = 2V$.

The information structure is as follows. At time 1, investor B_i observes $S_{B,i}$, and at time 2, investor A_i observes $S_{A,i}$ (and analogously for stock j). The terminal dividends on the stocks are given by

$$D_i = (S_{A,i} + S_{B,i})/2 + \epsilon_i, \tag{27}$$

$$D_j = (S_{A,j} + S_{B,j})/2 + \epsilon_j, \tag{28}$$

where $\text{cov}(S_{B,i}, S_{B,j}) = \text{cov}(S_{A,i}, S_{A,j}) > 0$. Thus the signals of the B investors are correlated with one another, as are the signals of the A investors. This can be thought of as reflecting the existence of common components across stocks i and j.[23]

The pricing of the two stocks proceeds as follows. First, note that because we have assumed that $H = 2V$, by Lemma 2, no information at all comes out for either stock at time 1—both $S_{B,i}$ and $S_{B,j}$ remain completely hidden at this time. Next, at time 2, we assume that the markets for the two stocks are momentarily segmented. That is, the auction mechanisms for the two are run separately. Thus the auction for stock i at time 2 only involves investors A_i and B_i. This implies that prices at time 2 are given exactly as in Lemmas 5–8, with the appropriate i and j subscripts.

However, right after the individual markets clear in segmented fashion at time 2, we insert another date, which we call time 2+, at which point we allow the arbitrageurs to look at prices in both markets simultaneously, in order to update their estimates. Thus for example, if $S_{B,i}$ is fully revealed in the stock i auction at time 2, it does not appear in the price of stock j at time 2, but it does immediately afterward, at time 2+.[24]

It should now be apparent how there can be contagion in this framework. Suppose that at time 2, there is no news about stock j—that is, $S_{A,j}$ stays just at its prior value of $(V + H)$—but there is bad firm-specific news about

---

[22] This sharper definition appears in many accounts of emerging markets crises, where the puzzle posed often goes something like this: Why is it that bad news that would seem to be specifically about the Russian economy not only devastates the Russian stock market but somehow leads to a drop in the Brazilian market as well? See, for example, Kodres and Pritsker (2002) for a recent treatment and references to related work.

[23] As will become clear, only the asumption that $\text{cov}(S_{B,i}, S_{B,j}) > 0$ is needed for our results. The value of $\text{cov}(S_{A,i}, S_{A,j})$ is not relevant; we just set it equal to $\text{cov}(S_{B,i}, S_{B,j})$ for symmetry.

[24] The momentary segmentation of the two markets before we allow full information sharing at time 2+ is just a modeling trick. We do it so as to keep the auctions for the two stocks separate, which greatly simplifies the analysis of equilibrium in terms of cutoff levels, etc. Indeed, it allows us to use all our earlier results from Lemmas 5–8 without any modification.

stock i, in the form of a low realization of $S_{A,i}$. Initially, in the segmented time 2 auctions, the price of stock j is unchanged. However, the price of stock i falls, and previously hidden information about $S_{B,i}$ is revealed, for exactly the same reasons as in the one-stock case. When we allow for information sharing at time 2+, this new information on $S_{B,i}$ causes investors in stock j to revise their estimates of $S_{B,j}$, given that $cov(S_{B,i}, S_{B,j}) > 0$. So, remarkably, the price of stock j moves even though there is absolutely no contemporaneous news about its own fundamentals. This in turn induces a correlation between the two stocks at a time when stock i is falling.

As a consequence, this setup also delivers increased comovements during market declines. To be specific, we define the price of stock i at time 2+ as $P_{i,2+}$, and the corresponding two-period returns:[25]

$$R_i = (P_{i,2+} - P_{i,0}), \tag{29}$$

$$R_M = (R_i + R_j)/2. \tag{30}$$

Here R_M is the return on a "market portfolio" that is an equal-weighted combination of stocks i and j. In addition, we define $\hat{\sigma}_{ij}$ to be the sample estimator of the covariance between the returns on stocks i and j:

$$\hat{\sigma}_{ij} = R_i R_j. \tag{31}$$

Note that now $\hat{\sigma}_{ij}$ is a random variable that depends on the realizations of both R_i and R_j.

We prove in the appendix that

Proposition 7. *The covariance estimator for stocks i and j covaries negatively with the return on the market portfolio*: $cov(\hat{\sigma}_{ij}, R_M) < 0$.

4. Empirical Implications

Throughout the article we have attempted to argue that our model does a parsimonious job of fitting a range of existing empirical facts about asymmetries in return distributions, large price movements, etc. But of course, any theory is more attractive if it also offers some novel, as-yet-untested predictions, and thereby puts itself at risk of being rejected in the data.

We believe that Proposition 5—which says that high trading volume should be associated with more negative skewness—may be particularly useful in this regard. Indeed, in work that was initiated after the first draft of this article was completed [Chen, Hong, and Stein (2001)], we conduct a series of cross-sectional tests that are motivated by Proposition 5. We begin by constructing

---

[25] Note that it is no longer meaningful to distinguish between one- and two-period returns, since given our simplifying assumption that $H = 2V$, prices do not change at time 1.

for individual firms a measure of the realized skewness of their daily returns, using six months of data at a time. We then try to forecast this skewness variable, using only prior information. For example, we would try to forecast the realized daily skewness for a given firm over the period January 1, 1999 to June 30, 1999, using information about that firm available prior to January 1. One of the variables that we use to do this forecast is the firm's detrended turnover from the previous six months (July 1, 1998–December 31, 1998, in this example). As it turns out, when turnover is high relative to trend, our estimates suggest that subsequent skewness is in fact more negative. Moreover, across a variety of specifications, the coefficients on turnover are strongly statistically significant, as well as economically meaningful.[26]

While the cross-sectional results in Chen, Hong, and Stein (2001) are supportive of the theory, we also document other pieces of evidence that, on the face of it, appear harder to reconcile with our model. Most notably, the average skewness of daily returns for individual stocks is positive, in marked contrast to the significant negative skewness in the returns of the market portfolio. Although our model admits both positive and negative values of skewness (depending on the heterogeneity parameter H), it would be a bit of a stretch to claim that it can explain why individual stocks are positively skewed, since there is no clear basis for assuming that there is less heterogeneity of opinion about individual stocks than about the market as a whole.[27]

In Chen, Hong, and Stein (2001), we appeal to factors outside of the current model to explain the positive skewness in individual stocks. In particular, we develop a "discretionary disclosure" hypothesis in which a manager sitting on a large chunk of good news releases it to the public immediately, but a manager sitting on bad news only allows this bad news to dribble out slowly. Such a pattern of information revelation will tend to impart a degree of positive skewness to individual stock returns, but will—if the managed news is idiosyncratic—have no effect on the skewness of the market portfolio. In support of this hypothesis, we show that positive skewness is most pronounced in small stocks, and, controlling for size, in stocks with low analyst coverage. To the extent that it is easier for the managers of small, low-coverage firms to conceal information from the market, this is exactly what one would predict based on the premise of discretionary disclosure.

Overall, this discussion suggests the following conclusions. On the one hand, the model in this article seems to do a good job of fitting the stylized

---

[26] A recent article by Dennis and Mayhew (2002), though it has a quite different motivation, also produces evidence that bears on our hypothesis. Dennis and Mayhew develop, for individual stocks, a measure of implied skewness based on options prices. They then regress this implied skewness measure against a variety of firm-level characteristics. One finding is that, controlling for size, options-implied negative skewness is more pronounced for high-turnover firms.

[27] Moreover, one might argue that our model predicts more negative skewness for individual stocks than for the market, to the extent that the former are harder to short, either due to direct transactions costs of shorting or to institutional restrictions.

facts about the aggregate market sketched in the introduction. It also seems to be helpful in explaining the cross section of skewness in individual stocks. On the other hand, the model, as it currently stands, does not provide a good story for the average level of skewness in individual stocks. If it is to do so, it will likely have to be extended, perhaps along the lines of our informal discretionary disclosure hypothesis.

5. Related Work

In this section we discuss the link between our theory and previous work that also focuses on either large price movements or return asymmetries. We divide this other work into three broad categories: (1) rational models with incomplete information aggregation; (2) volatility feedback models; and (3) behavioral stories.

5.1 Rational models with incomplete information aggregation

One important class of theories shows how there can be large movements in asset prices in the absence of external news about fundamentals, even when all market participants are fully rational. Notable articles include Grossman (1988), Gennotte and Leland (1990), Jacklin, Kleidon, and Pfleiderer (1992), and Romer (1993).[28] All these articles share a common theme: investors are initially imperfectly informed about some important variable, which is not revealed to them in prices. However, the process of trading may eventually cause this information to come out, at which time prices can change sharply, even if no external news has arrived.

Although our model incorporates some less than fully rational agents, it clearly draws heavily on the basic insights from this earlier work, particularly Romer (1993). Romer's model has the feature that traders start out not knowing the precision of each others' information. As shocks arrive, they update their estimates of this precision, which can be given the interpretation that each trader is learning about the elasticity of demand of other traders. This is broadly analogous to our idea of the arbitrageurs learning about the extent of "buying support" from investor B as investor A bails out of the market.

In spite of the similarities, however, there is one crucial distinction between our model and these others: ours is fundamentally asymmetric, producing larger downward price movements than upward ones. In contrast, all the logic in Romer (1993) is totally symmetric, so large up and down moves

---

[28] See Kleidon (1995) for a detailed review of this branch of the literature and further references.

are equally likely.[29] Somewhat more subtly, while the models of Grossman (1988), Gennotte and Leland (1990), and Jacklin, Kleidon, and Pfleiderer (1992) have been used specifically to explain the stock market crash of October 1987, this argument entails the additional assumption that market participants systematically underestimated the extent of portfolio insurance that was in place prior to the crash. While such a "one-bad-draw" type of assumption may be a perfectly reasonable way to rationalize a single given event, it is hard to use these sorts of models if the goal is to explain a pervasive tendency of markets to melt down, rather than melt up. In other words, if one runs these models over repeatedly, on average traders should be just as likely to overestimate as to underestimate the extent of portfolio insurance, so the long-run distribution should show just as many big up moves as big down moves.[30]

5.2 Volatility feedback models
Unlike the incomplete information aggregation theories, the volatility feedback literature, which includes Pindyck (1984), French, Schwert, and Stambaugh (1987), and Campbell and Hentschel (1992), is all about asymmetries in returns. The basic idea is a simple one. When a large piece of news arrives, this signals that market volatility has gone up. Assuming that market volatility commands a risk premium, the positive effect of a large piece of good news is damped, as the increased risk premium effect partially offsets the direct good news effect. Conversely the negative effect of a large piece of bad news is amplified. The result is that even if the process driving news is symmetric, returns—particularly the returns on the market portfolio—will be negatively skewed.

One important difference between our approach and the volatility feedback models is that they are set in a representative agent framework, and hence are completely silent on trading volume, while our focus on differences of opinion leads very naturally to a linkage between volume and the conditional intensity of negative skewness. Moreover, from the perspective of our interest in crashes, a weakness of the volatility feedback class of models is that, like any rational model with perfect information, they require very large doses of external public news if they are to generate very large price movements.[31]

---

[29] A recent article that extends Romer (1993) is Cao, Coval, and Hirshleifer (2002). Their model can produce certain conditional patterns in skewness—that is, negative skewness after price increases, and positive skewness after price declines. However, it appears that, like in Romer's model, there is no unconditional skewness, and hence no prediction that large down moves are ex ante more likely than large up moves.

[30] A similar observation can be made about other, more microstructure-oriented accounts of the October 1987 crash, including those in the Brady Report (1988) or Greenwald and Stein (1991). These accounts emphasize how unusually large selling volume overwhelmed market-making capacity, thereby exacerbating the decline in prices. But their inherently symmetric logic would seem to suggest that if there were ever comparable buying volume, one should get an equal-sized increase in prices.

[31] The same can be said of the "leverage effects" analyzed by Black (1976), Christie (1982), and Schwert (1989): they can clearly create asymmetries, but they cannot deliver price changes that are out of proportion to contemporaneous public news.

Thus these models may be more helpful in thinking about skewness in "typical" stock returns, as opposed to rare tail events; certainly this has been the focus of the empirical literature that has adopted the volatility feeback approach.[32]

So to give an oversimplified summary, the incomplete information aggregation models and the volatility feedback models each get at some of the elements of our definition of a crash. But none get at all the elements simultaneously. The incomplete information aggregation models are good for producing big price movements without big news, but not for delivering asymmetries. Conversely, the volatility feedback models generate asymmetries, but they do not leave open a role for trading volume. Nor is it clear that they are helpful for thinking about really dramatic price changes, particularly if these price changes occur in the absence of equally dramatic news.

5.3 Behavioral stories

It is often argued informally that stock market crashes should be thought of as evidence against traditional, fully rational models of asset pricing. For example, Shleifer and Summers (1990:19) write: "the stock in the efficient markets hypothesis . . . crashed along with the rest of the market on October 19, 1987." While this sentiment may well be on target, it strikes us as somewhat ironic that the behavioral finance literature has not really made much progress in understanding crashes; indeed, in our view, not nearly as much as the above-discussed line of work on incomplete information aggregation.

A standard behavioral interpretation of a market crash is that it represents a sudden, radical shift in investor sentiment; in the words of Shiller (1989:1), a crash is a time when "the investing public en masse capriciously changes its mind." But, as with the more rational theories, this explanation again leaves unanswered the question of asymmetries: why is it that the biggest capricious changes in sentiment are negative, rather than positive changes? Perhaps one might argue that fear and panic are more powerful emotions than optimism and euphoria, but this strikes us as an unsatisfying rationalization.[33]

In contrast, one of the key selling points of our theory is that one can start with a symmetric driving process for investors' beliefs and still—with the help of the short-sales constraint—generate asymmetries in returns. Certainly one can, in the spirit of Shiller, think of the signals in our model as containing an element of capricious investor sentiment. But now we can say that even if investor *A* is as likely to become overoptimistic as overpessimistic at time 2, the market is more likely to melt down than to melt up at this time.

---

[32] A related point, due to Poterba and Summers (1986), is that the quantitative significance of the volatility feedback effect is likely to be small, since shocks to market volatility are not very persistent.

[33] As an alternative to investor sentiment stories, there are also rational bubble models of the sort described by Blanchard and Watson (1982). The popping of a stochastic bubble can be interpreted as a market crash, and it satisfies our criteria in terms of being both a big move in the absence of news, as well as an inherently asymmetric phenomenon. However, bubble models have not fared well empirically [West (1988), Flood and Hodrick (1990)]. Moreover, they have the unattractive feature that the crash, when it occurs, is based on the realization of an extrinsic "sunspot," and hence cannot be explained within the context of the model.

6. Conclusion

The model in this article is meant to capture a simple—and we believe important—piece of intuition about the effect of short-sales constraints on stock prices: the private information of those relatively bearish investors who are initially sidelined by the short-sales constraint is more likely to be flushed out through the trading process when the market is falling, as opposed to rising. As we have argued, this mechanism can help shed light on a variety of stylized facts, including: (1) large movements in prices that are not accompanied by significant contemporaneous news about fundamentals; (2) negative skewness in the distribution of market returns; and (3) increased correlation among stocks in a falling market.

In addition to rationalizing these existing pieces of evidence, the model also makes a distinctive new out-of-sample prediction, that negative skewness will be most pronounced conditional on high trading volume. Recent work reported in Chen, Hong, and Stein (2001) lends support to this prediction. At the same time, this empirical work also makes it clear that the current version of the model cannot easily come to grips with all of the relevant facts about asymmetries in returns: while the model seems to do a good job of explaining cross-sectional variation in skewness at the firm level (skewness is more negative in firms that have recently experienced increases in trading volume) it has a hard time with the fact that on average, skewness in individual-firm returns is positive. In order to square with this last fact, the model will most likely have to be extended, possibly by incorporating the notion that managers have some control over the rate at which information is disclosed to the market, and that they are less inclined to let bad news get out promptly.

This article can be seen as part of a recent resurgence of theoretical and empirical interest in the general topic of how short-sales constraints shape stock prices. A prominent theme in much of this new work is the interplay of short-sales constraints, less than fully rational investors, and more rational arbitrageurs.[34] This work is also beginning to suggest that short-sales constraints may play a bigger role than one might have guessed based on just the direct transactions costs associated with shorting. Although these direct transactions costs can be substantial in a minority of cases [D'Avolio (2002), Geczy, Musto, and Reed (2002), Lamont and Jones (2002)], shorting is also meaningfully constrained by other factors, including institutional restrictions (as in the case of many mutual funds), and possibly some less-well-understood behavioral biases. Unfortunately our understanding of these nonstandard constraints is still very sketchy. There remains much to be done, both in terms of developing a fuller understanding of why so many investors behave as if they were facing prohibitive shorting costs, and of exploring the consequences of such behavior for stock prices.

---

[34] See, for example, Chen, Hong, and Stein (2002), D'Avolio (2002), Geczy, Musto, and Reed (2002), Lamont and Jones (2002), Ofek and Richardson (2003), and Diether, Malloy, and Scherbina (2002).

Appendix A: Proofs of Lemmas

Proof of Lemma 1.   Since S_B is uniform on $[0, 2V]$, the risk-neutral arbitrageurs' forecast of S_B given that investor B has not submitted an order at trial price p_1 is given by

$$E[S_B|S_B \leq p_1] = p_1/2; \tag{A.1}$$

hence the risk-neutral arbitrageurs' estimate of the terminal value of the asset is

$$E[D|S_B \leq p_1] = (V+H)/2 + p_1/4. \tag{A.2}$$

For any $p_1 > S_B^*$, it is easy to show that $E[D|S_B \leq p_1] - p_1 < 0$ and so the auctioneer lowers p_1 as long as investor B does not submit an order. But since $S_B > S_B^*$, then it follows that S_B will be revealed for a low enough p_1. Hence,

$$P_1 = (V+H)/2 + S_B/2 \tag{A.3}$$

will be the equilibrium price.

Proof of Lemma 2.   Suppose $S_B \leq S_B^*$. Then for all trial prices p_1, where $p_1 \geq S_B^*$, it follows that $S_B - p_1 \leq 0$ and investor B never reveals his information. At $p_1 = S_B^*$, the arbitrageurs' expectation of the stock's terminal value,

$$E[D|p_1 = S_B^*] = (V+H)/2 + S_B^*/4, \tag{A.4}$$

equals the prevailing trial price and so the market clears at

$$P_1 = (V+H)/2 + S_B^*/4 = 2(V+H)/3, \tag{A.5}$$

before investor B ever gets into the market. The ex ante probability of winding up in the pooling equilibrium is just

$$S_B^*/(2V) = (V+H)/(3V), \tag{A.6}$$

the probability that $S_B \leq S_B^*$.

Proof of Lemma 3.   Since S_A is uniform on $[H, 2V+H]$, the risk-neutral arbitrageurs' forecast of S_A given that investor A has not submitted an order at trial price p_2 is

$$E[S_A|S_A \leq p_2] = (H+p_2)/2; \tag{A.7}$$

hence the risk-neutral arbitrageurs' estimate of the terminal value of the asset is

$$E[D|S_A \leq p_2] = (H+p_2)/4 + S_B/2. \tag{A.8}$$

For any $p_2 > S_A^*$, it is easy to show that $E[D|S_A \leq p_2] - p_2 < 0$ and so the auctioneer lowers p_2 as long as investor A does not submit an order. But since $S_A > S_A^*$, then it follows that S_A will be revealed for a low enough p_2. Hence,

$$P_2 = S_A/2 + S_B/2 \tag{A.9}$$

will be the equilibrium price.

Proof of Lemma 4. Suppose $S_A \leq S_A^*$. Then for all trial prices p_2, where $p_2 \geq S_A^*$, it follows that $S_A - p_2 \leq 0$ and investor A never reveals his information. At $p_2 = S_A^*$, the arbitrageurs' expectation of the stock's terminal value,

$$E[D|p_2 = S_A^*] = (H + S_A^*)/4 + S_B/2, \tag{A.10}$$

equals the prevailing trial price and so the market clears at

$$P_2 = (H + S_A^*)/4 + S_B/2 = (2S_B + H)/3 \tag{A.11}$$

before investor A ever gets into the market.

Proof of Lemma 5. We know that no new information regarding investor B's valuation arrives at $t = 2$ and (from Lemma 2) S_B was hidden at equilibrium price

$$P_1 = (V + H)/2 + S_B^*/4. \tag{A.12}$$

If $S_A > (V + H)$, then it follows that $S_A - p_2 > 0$ for some $p_2 > P_1$ and hence S_A will be revealed for some $p_2 > P_1$. So,

$$P_2 = S_A/2 + S_B^*/4 = S_A/2 + (V + H)/6 \tag{A.13}$$

will be the equilibrium price, and the market clears before any new information on S_B can be revealed.

Proof of Lemma 6. Since $S_B \leq S_B^{**}$, where $S_B^{**} = 2S_A/3$, it follows that S_A will be revealed before S_B during the auction. Given that S_A has been revealed, the risk-neutral arbitrageurs' estimate of the terminal value of the asset at trial price p_2 is

$$E[D|S_B \leq p_2] = S_A/2 + p_2/4. \tag{A.14}$$

Since $S_B \leq S_B^{**}$, then for all trial prices p_2, where $p_2 \geq S_B^{**}$, it follows that $S_B - p_2 \leq 0$ and investor B never reveals his information. At $p_2 = S_B^{**}$, the arbitrageurs' expectation of the stock's terminal value,

$$E[D|p_2 = S_B^{**}] = S_A/2 + S_B^{**}/4, \tag{A.15}$$

equals the prevailing trial price and so the market clears at

$$P_2 = S_A/2 + S_B^{**}/4 = 2S_A/3 \tag{A.16}$$

before investor B ever gets into the market.

Proof of Lemma 7. It is easy to verify that if $S_A \leq S_A^*$ and $S_B > H$ (given that $S_B < S_B^*$), then $S_A < S_B$. Hence S_B will be revealed before S_A. Given that S_B has been revealed, the risk-neutral arbitrageurs' estimate of the terminal value of the asset at trial price p_2 is

$$E[D|S_A < p_2] = (H + p_2)/4 + S_B/2. \tag{A.17}$$

Since $S_A \leq S_A^*$, then for all trial prices p_2, where $p_2 \geq S_A^*$, it follows that $S_A - p_2 \leq 0$ and investor A never reveals his information. At $p_2 = S_A^*$, the arbitrageurs' expectation of the stock's terminal value,

$$E[D|p_2 = S_A^*] = (H + S_A^*)/4 + S_B/2, \tag{A.18}$$

equals the prevailing trial price and so the market clears at

$$P_2 = (H + S_A^*)/4 + S_B/2 = (2S_B + H)/3 \tag{A.19}$$

before investor A ever gets into the market.

Proof of Lemma 8. It is easy to verify from Figure 1 that the only parameter region that remains is the case where $2(V+H)/3 > S_B > 2H/3$ and for a fixed realization of S_B, $S_A \in ((2S_B + H)/3, 3S_B/2)$. In this parameter region, S_A and S_B are not too far apart and S_B is not too small, so both values are revealed in the auction process. More precisely, the conditions required for either one to remain hidden—as established in the previous lemmas—cannot be established. So,

$$P_2 = S_A/2 + S_B/2 \tag{A.20}$$

is the equilibrium price.

Proof of Lemma 9. Let $P_0 = V + H/2$. Then

$$E[R_1^3] = E[R_1^3 | S_B > S_B^*] \Pr(S_B > S_B^*) + E[R_1^3 | S_B \le S_B^*] \Pr(S_B \le S_B^*). \tag{A.21}$$

Then from the fact that S_B is uniform on $[0, 2V]$,

$$E[R_1^3] = \frac{1}{2V} \left[\int_{S_B^*}^{2V} \left(\frac{V+H+y}{2} - P_0 \right)^3 dy + \int_0^{S_B^*} \left(S_B^* - P_0 \right)^3 dy \right]. \tag{A.22}$$

It can be shown that

$$E[R_1^3] = -\frac{(H - 2V)(H + V)^3}{648V}. \tag{A.23}$$

Since $H \le 2V$, it follows that $\forall H$ and V, $E[R_1^3] \ge 0$.

Proof of Lemma 10. First, we calculate $E[R_2^3]$ conditional on being in Case 1, in which S_B is revealed at $t = 1$, that is, $S_B > S_B^*$. For convenience, let $x = S_A$ and $y = S_B$. Under Case 1, the price is simply $P_1 = (V + H)/2 + y/2$. Then

$$E[R_2^3 \,|\, \text{Case } 1] = \frac{1}{(2V - S_B^*)} \frac{1}{2V}$$
$$\times \int_{S_B^*}^{2V} \left[\int_H^{\frac{2y+H}{3}} \left(\frac{2y+H}{3} - P_1 \right)^3 dx + \int_{\frac{2y+H}{3}}^{2V+H} \left(\frac{x+y}{2} - P_1 \right)^3 dx \right] dy. \tag{A.24}$$

Since S_A and S_B are independent and uniformly distributed, we have that

$$E[R_2^3 \,|\, \text{Case } 1] = -\frac{(H - 2V)^3 (121H + 208V)}{262440V}. \tag{A.25}$$

Here, since $H \le 2V$, it follows that $E[R_2^3 | \text{Case } 1] \ge 0$.

Now, we calculate $E[R_2^3]$ conditional on being in Case 2, in which S_B is hidden at $t = 1$, that is, $S_B \le S_B^*$. Under Case 2, the price is simply $P_1 = S_B^*$. We will calculate

$$E[R_2^3 \,|\, \text{Case } 2] = \frac{1}{S_B^*} \frac{1}{2V} \sum_{i=1}^3 E^i \tag{A.26}$$

where E^1 is given by

$$E^1 = \int_0^{2H/3} \left[\int_H^{V+H} \left(\frac{x}{2} + \frac{2x/3}{4} - P_1 \right)^3 dx + \int_{V+H}^{2V+H} \left(\frac{x}{2} + \frac{V+H}{6} - P_1 \right)^3 dx \right] dy; \quad \text{(A.27)}$$

E^2 is given by

$$E^2 = \int_{2H/3}^H \left[\int_H^{\frac{3y}{2}} \left(\frac{x+y}{2} - P_1 \right)^3 dx + \int_{\frac{3y}{2}}^{V+H} \left(\frac{x}{2} + \frac{2x/3}{4} - P_1 \right)^3 dx \right.$$

$$\left. + \int_{V+H}^{2V+H} \left(\frac{x}{2} + \frac{V+H}{6} - P_1 \right)^3 dx \right] dy; \quad \text{(A.28)}$$

and E^3 is given by

$$E^3 = \int_H^{S_B^*} \left[\int_H^{\frac{2y+H}{3}} \left(\frac{2y+H}{3} - P_1 \right)^3 dx + \int_{\frac{2y+H}{3}}^{\frac{3y}{2}} \left(\frac{x+y}{2} - P_1 \right)^3 dx \right.$$

$$\left. + \int_{\frac{3y}{2}}^{V+H} \left(\frac{x}{2} + \frac{2x/3}{4} - P_1 \right)^3 dx + \int_{V+H}^{2V+H} \left(\frac{x}{2} + \frac{V+H}{6} - P_1 \right)^3 dx \right] dy. \quad \text{(A.29)}$$

Since S_A and S_B are independent and uniformly distributed, we have that

$$E[R_2^3 \,|\, \text{Case 2}] = -\frac{7H^5 - 70H^4V - 935H^3V^2 + 5515H^2V^3 + 9065HV^4 + 2935V^5}{1049760V(H+V)}. \quad \text{(A.30)}$$

Without loss of generality, let $H = \alpha V$, where $\alpha \in [0, 2]$, that is, $\alpha = H/V$. Let

$$f(\alpha) = \frac{E[R_2^3 \,|\, \text{Case 2}]}{V^3}. \quad \text{(A.31)}$$

Note that $f(0) = -587/209952$ and $f(2) = -143/1296$, where $f(0) < f(2)$. Furthermore,

$$f'(\alpha) = -\frac{6130 + 11030\alpha + 2710\alpha^2 - 2150\alpha^3 - 175\alpha^4 + 28\alpha^5}{1049760(1+\alpha)^2}. \quad \text{(A.32)}$$

It is easy to show that $\forall \alpha, f'(\alpha) < 0$. Note also that $\max f'(\alpha) = -41/19440$ and $\text{argmax}\, f'(\alpha) = 2$. Hence, $\forall H \in [0, 2V], E[R_2^3 \,|\, \text{Case 2}] < 0$; moreover, $E[R_2^3 \,|\, \text{Case 2}]$ decreases monotonically with H/V.

Since the probability of being in Case 1 is $(2V - S_B^*)/2V$ and the probability of being in Case 2 is $S_B^*/2V$, it follows that

$$E[R_2^3] = E[R_2^3 \,|\, \text{Case 1}] \frac{(2V - S_B^*)}{2V} + E[R_2^3 \,|\, \text{Case 2}] \frac{S_B^*}{2V}. \quad \text{(A.33)}$$

Simplifying, we have

$$E[R_2^3] = \frac{53H^5 - 330H^4V + 655H^3V^2 - 115H^2V^3 - 3105HV^4 + 1153V^5}{349920V^2}. \quad \text{(A.34)}$$

Let

$$g(\alpha) = E[R_2^3]/V^3. \quad \text{(A.35)}$$

Note that $g(0) = 1153/349920$ and $g(2) = -143/12960$. Furthermore,

$$g'(\alpha) = \frac{-3105 - 230\alpha + 1965\alpha^2 - 1320\alpha^3 + 265\alpha^4}{349920}. \tag{A.36}$$

It is easy to show that $\forall \alpha$, $g'(\alpha) < 0$ since $\max g'(\alpha) = -5/864$ and $\operatorname{argmax} g'(\alpha) = 2$. Moreover, $E[R_2^3] = 0$ for $H/V \approx .3756$. Hence, $E[R_2^3]$ monotonically decreases with H/V and is negative for $H/V > .3756$.

Proof of Lemma 11. We have already observed that for $H > V$, there is no trade at time 1. Thus total trading volume will be equivalent to the long position of whichever investor (A or B) has the higher valuation at time 2, that is, volume will be proportional to $(S_A - P_2)$ if $S_A > S_B$ and to $(S_B - P_2)$ if $S_B > S_A$.

Next, observe that when $S_A > S_B$, we can rewrite $P_2 = S_A/2 + \widehat{S}_B/2$, where $\widehat{S}_B/2$ is the arbitrageurs' estimate of S_B (whether it is fully revealed or not). Thus when $S_A > S_B$, volume is proportional to $(S_A/2 - \widehat{S}_B/2)$. Conversely, when $S_B > S_A$, volume is proportional to $(S_B/2 - \widehat{S}_A/2)$.

Applying the law of iterated expectations, when $S_A > S_B$, ex ante expected volume is proportional to $E[S_A/2 - \widehat{S}_B/2] = E[S_A/2 - S_B/2]$. Similarly, when $S_B > S_A$, ex ante expected volume is proportional to $E[S_B/2 - \widehat{S}_A/2] = E[S_B/2 - S_A/2]$. Combining terms, we can write overall ex ante expected volume as $E[|S_A/2 - S_B/2|]$ as in the text.

Now it is straightforward to evaluate $E[|S_A/2 - S_B/2|]$ by integrating over the distributions of S_A and S_B. Performing this integration, we find that $E[|S_A/2 - S_B/2|] = H/4$. So ex ante expected trading volume equals $kH/4$, as stated in the lemma.

Appendix B: Proofs of Propositions

Proof of Proposition 1. Since S_B and S_A are symmetrically distributed (with the same variance) and are fully revealed at time 1 and 2, respectively, the fact that prices are efficient means that returns are symmetrically distributed and homoscedastic.

Proof of Proposition 2. This follows easily from the calculations provided in Section 2.3.1 of the text.

Proof of Proposition 3. From Equations (A.23) and (A.34), we have that

$$E[R_s^3] = \frac{2233V^5 - 405HV^4 + 1505H^2V^3 + 115H^3V^2 - 870H^4V + 53H^5}{699840V^2}. \tag{A.37}$$

Then let

$$h(\alpha) = \frac{E[R_s^3]}{V^3}. \tag{A.38}$$

Note that $h(0) = 223/699840$, $h(2) = -143/25920$, and $h(1.691) = 0$. Using similar arguments to those in Lemma 10, it is not hard to show that for $H/V > 1.691$, $E[R_s^3] < 0$.

Proof of Proposition 4. We now calculate the medium-horizon skewness $E[R_m^3]$. Using the same arguments as in Lemmas 9 and 10, we have

$$E[R_m^3] = \frac{-8V^5 + 145HV^4 - 30H^2V^3 + 40H^3V^2 - 50H^4V + H^5}{51840\sqrt{2}}. \tag{A.39}$$

From Equations (A.37) and (A.39), we can define the difference between the short-horizon and the medium-horizon skewness as

$$\Delta = \frac{E[R_s^3] - E[R_m^3]}{V^3}. \tag{A.40}$$

It is not hard to show that $\forall \, \alpha > 1.834$, $\Delta < 0$. So for sufficiently large values of H/V, there is less negative skewness in medium-horizon returns than in short-horizon returns.

Proof of Proposition 6.   We will present the proof for $\hat{\rho}_{ij}$; the proof for $\hat{\sigma}_{ij}$ is a simpler variation on the same argument. Recall our definition of $\hat{\rho}_{ij}$:

$$\hat{\rho}_{ij} = \frac{E[R_i R_j | R_M]}{\sqrt{E[R_i^2 | R_M] E[R_j^2 | R_M]}}. \tag{A.41}$$

Since $R_i = R_M + Z_i$, it follows that

$$\hat{\rho}_{ij} = \frac{R_M^2}{R_M^2 + \sigma_Z^2}. \tag{A.42}$$

where σ_Z^2 is the variance of the Z_i's. Then observe that

$$\frac{1}{\hat{\rho}_{ij} - 1} = -\frac{R_M^2}{\sigma_Z^2} - 1. \tag{A.43}$$

So we have

$$E\left[\frac{1}{\hat{\rho}_{ij} - 1} R_M\right] = E\left[-\frac{R_M^3}{\sigma_Z^2} - R_M\right] = -\frac{E[R_M^3]}{\sigma_Z^2} > 0, \tag{A.44}$$

since $E[R_M^3] < 0$ by assumption. It follows then that

$$E[\hat{\rho}_{ij} R_M] < 0, \tag{A.45}$$

since

$$\frac{\partial}{\partial \hat{\rho}_{ij}} \left(\frac{1}{\hat{\rho}_{ij} - 1}\right) < 0. \tag{A.46}$$

Proof of Proposition 7.   We first define some simplifying notations. Let

$$\widehat{S}_{B,k} = E_{2+}[S_{B,k}] \tag{A.47}$$

be the conditional expectation of $S_{B,k}$ at time $t = 2+$ for $k = i, j$. Then define the operator δ, applied to a random variable X, as

$$\delta X = X - E_0[X]. \tag{A.48}$$

We can then rewrite

$$R_i = \frac{1}{2}(\delta S_{A,i} + \delta \widehat{S}_{B,i}),$$

$$R_j = \frac{1}{2}(\delta S_{A,j} + \delta \widehat{S}_{B,j}), \tag{A.49}$$

$$R_M = \frac{1}{4}(\delta S_{A,i} + \delta S_{A,j} + \delta \widehat{S}_{B,i} + \delta \widehat{S}_{B,j}).$$

Using the expressions for R_i, R_j, and R_M given in Equation (A.49), we can expand $E[R_iR_jR_M]$. Dropping terms that are equal to zero, we have

$$E[R_iR_jR_M] = \frac{1}{16}E[\delta S_{A,j}(\delta\widehat{S}_{B,i})^2 + \delta S_{A,i}(\delta\widehat{S}_{B,j})^2$$
$$+ 2\delta S_{A,i}\delta\widehat{S}_{B,i}\delta\widehat{S}_{B,j} + 2\delta S_{A,j}\delta\widehat{S}_{B,i}\delta\widehat{S}_{B,j}]. \tag{A.50}$$

We already know from the one-asset version of our model that $E[\delta S_{A,k}(\delta\widehat{S}_{B,k})^2] < 0$ for $k = i, j$. Given the further assumptions that $\text{cov}(S_{B,i}, S_{B,j}) > 0$, and that the S_A's and S_B's are independent, it must be that

$$\delta\widehat{S}_{B,i} = k\delta\widehat{S}_{B,j} + \eta_i \tag{A.51}$$

(and vice versa), where $k > 0$ and η_i is independent of $S_{A,i}$ and $S_{A,j}$. It then follows that the expectation of each of the terms of Equation (A.50) is negative, so $E[R_iR_jR_M] < 0$.

References

Almazan, A., K. C. Brown, M. Carlson, and D. A. Chapman, 2001, "Why Constrain Your Mutual Fund Manager?," working paper, University of Texas.

Bakshi, G., C. Cao, and Z. Chen, 1997, "Empirical Performance of Alternative Option Pricing Models," *Journal of Finance*, 52, 2003–2049.

Bates, D., 2001, "Post-'87 Crash Fears in the S&P 500 Futures Options Market," *Journal of Financial Econometrics*, 94, 181–238.

Bekaert, G., and G. Wu, 2000, "Asymmetric Volatility and Risk in Equity Markets," *Review of Financial Studies*, 13, 1–42.

Black, F., 1976, "Studies of Stock Price Volatility Changes," *Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economical Statistics Section*, 177–181.

Blanchard, O. J., and M. W. Watson, 1982, "Bubbles, Rational Expectations, and Financial Markets," in Paul Wachtel (ed.), *Crises in Economic and Financial Structure*, Lexington Books, Lexington, MA.

Brady Report, 1988, Presidential Task Force on Market Mechanisms, U.S. Government Printing Office, Washington, DC.

Braun, P. A., D. B. Nelson, and A. M. Sunier, 1995 "Good News, Bad News, Volatility, and Betas," *Journal of Finance*, 50, 1575–1603.

Campbell, J. Y., and L. Hentschel, 1992, "No News Is Good News: An Asymmetric Model of Changing Volatility in Stock Returns," *Journal of Financial Economics*, 31, 281–318.

Cao, H. H., J. D. Coval, and D. Hirshleifer, 2002, "Sidelined Investors, Trading Generated News and Security Returns," *Review of Financial Studies*, 15, 615–648.

Caplin, A., and J. Leahy, 1994, "Business as Usual, Market Crashes, and Wisdom After the Fact," *American Economic Review*, 84, 548–565.

Chen, J., H. Hong, and J. C. Stein, 2001, "Forecasting Crashes: Trading Volume, Past Returns, and Conditional Skewness in Stock Prices," *Journal of Financial Economics*, 61, 345–381.

Chen, J., H. Hong, and J. C. Stein, 2002, "Breadth of Ownership and Stock Returns," *Journal of Financial Economics*, 66, 171–205.

Christie, A. A., 1982, "The Stochastic Behavior of Common Stock Variances—Value, Leverage and Interest Rate Effects," *Journal of Financial Economics*, 10, 407–432.

Cutler, D. M., J. M. Poterba, and L. H. Summers, 1989, "What Moves Stock Prices?," *Journal of Portfolio Management*, 15, 4–12.

D'Avolio, G., 2002, "The Market for Borrowing Stock," *Journal of Financial Economics*, 66, 271–300.

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.

Dennis, P., and S. Mayhew, 2002, "Risk-Neutral Skewness: Evidence from Stock Options," forthcoming in *Journal of Financial and Quantitative Analysis*.

Derman, E., 1999, "Regimes of Volatility: Some Observations on the Variation of S&P 500 Implied Volatilities," *Goldman Sachs Quantitative Strategies Research Notes*.

Diamond, D. W., and R. E. Verrechia, 1987, "Constraints on Short-Selling and Asset Price Adjustment to Private Information," *Journal of Financial Economics*, 18, 277–311.

Diether, K. B., C. J. Malloy, and A. Scherbina, 2002, "Differences of Opinion and the Cross-Section of Stock Returns," *Journal of Finance*, 57, 2113–2141.

Domowitz, I., and A. Madhavan, 1998, "Open Sesame: Alternative Open Algorithms in Securities Markets," working paper, USC Business School.

Duffee, G. R., 1995a, "Asymmetric Cross-Sectional Dispersion in Stock Returns: Evidence and Implications," working paper, Board of Governors.

Duffee, G. R., 1995b, "Stock Returns and Volatility: A Firm-Level Analysis," *Journal of Financial Economics*, 37, 399–420.

Dumas, B., J. Fleming, and R. E. Whaley, 1998, "Implied Volatility Functions: Empirical Tests," *Journal of Finance*, 53, 2059–2106.

Engle, R. F., and V. K. Ng, 1993, "Measuring and Testing the Impact of News on Volatility," *Journal of Finance*, 48, 1749–1778.

Flood, R. P., and R. J. Hodrick, 1990, "On Testing for Speculative Bubbles," *Journal of Economic Perspectives*, 4, 85–101.

French, K. R., and R. Roll, 1986, "Stock Return Variances: The Arrival of Information and the Reaction of Traders," *Journal of Financial Economics*, 17, 5–26.

French, K. R., G. W. Schwert, and R. F. Stambaugh, 1987, "Expected Stock Returns and Volatility," *Journal of Financial Economics*, 19, 3–29.

Genotte, G., and H. Leland, 1990, "Market Liquidity, Hedging and Crashes," *American Economic Review*, 80, 999–1021.

Geczy, C. C., D. K. Musto, and A. V. Reed, 2002, "Stocks are Special Too: An Analysis of the Equity Lending Market," *Journal of Financial Economics*, 66, 241–269.

Glosten, L., R. Jagannathan, and D. E. Runkle, 1993, "On the Relation Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks," *Journal of Finance*, 48, 1779–1801.

Greenwald, B. C., and J. C. Stein, 1991, "Transactional Risk, Market Crashes, and the Role of Circuit Breakers," *Journal of Business*, 64, 443–462.

Grossman, S. J., 1988, "An Analysis of the Implications for Stock and Futures Price Volatility of Program Trading and Dynamic Hedging Strategies," *Journal of Business*, 61, 275–298.

Harris, M., and A. Raviv, 1993, "Differences of Opinion Make a Horse Race," *Review of Financial Studies*, 6, 473–506.

Harrison, J. M., and D. M. Kreps, 1978, "Speculative Investor Behavior in a Stock Market with Heterogeneous Expectations," *Quarterly Journal of Economics*, 93, 323–336.

Hong, H., and J. C. Stein, 1999, "A Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets," *Journal of Finance*, 54, 2143–2184.

Jacklin, C. J., A. W. Kleidon, and P. Pfleiderer, 1992, "Underestimation of Portfolio Insurance and the Crash of October 1987," *Review of Financial Studies*, 5, 35–63.

Kandel, E., and N. D. Pearson, 1995, "Differential Interpretation of Public Signals and Trade in Speculative Markets," *Journal of Political Economy*, 103, 831–872.

Kelly, M., 1994, "Correlation: Stock Answer," *Risk*, 7, 40–43.

Kleidon, A. W., 1995, "Stock Market Crashes," in R. A. Jarrow, V. Maksimovic, and W. T. Ziemba (eds.), *Handbooks in OR and MS*, Vol. 9.

Kodres, L. E., and M. Pritsker, 2002, "A Rational Expectations Model of Financial Contagion," *Journal of Finance*, 57, 769–799.

Koski, J. L., and J. Pontiff, 1999, "How Are Derivatives Used? Evidence from the Mutual Fund Industry," *Journal of Finance*, 54, 791–816.

Lamont, O. A., and C. M. Jones, 2002, "Short Sale Constraints and Stock Returns," *Journal of Financial Economics*, 66, 207–239.

Nelson, D., 1991, "Conditional Heteroskedasticity in Asset Returns: A New Approach," *Econometrica*, 59, 347–370.

Odean, T., 1998, "Volume, Volatility, Price and Profit When all Traders Are Above Average," *Journal of Finance*, 53, 1887–1934.

Ofek, E., and M. Richardson, 2003, "DotCom Mania: The Rise and Fall of Internet Stock Prices," forthcoming in *Journal of Finance*.

Pindyck, R. S., 1984, "Risk, Inflation, and the Stock Market," *American Economic Review*, 74, 334–351.

Poterba, J. M., and L. H. Summers, 1986, "The Persistence of Volatility and Stock Market Fluctuations," *American Economic Review*, 76, 1142–1151.

Roll, R., 1984, "Orange Juice and Weather," *American Economic Review*, 74, 861–880.

Roll, R., 1988, "R^2," *Journal of Finance*, 43, 541–566.

Romer, D., 1993, "Rational Asset-Price Movements Without News," *American Economic Review*, 83, 1112–1130.

Schwert, G. W., 1989, "Why Does Stock Market Volatility Change Over Time?," *Journal of Finance*, 44, 1115–1153.

Shleifer, A., and L. H. Summers, 1990, "The Noise Trader Approach to Finance," *Journal of Economic Perspectives*, 4, 19–33.

Shiller, R. J., 1989, *Market Volatility*, MIT Press, Cambridge, MA.

Varian, H. R., 1989, "Differences of Opinion in Financial Markets," in C. C. Stone (ed.), *Financial Risk: Theory, Evidence and Implications: Proceedings of the 11th Annual Economic Policy Conference of the Federal Reserve Bank of St. Louis*, Kluwer Academic, Boston.

West, K. D., 1988, "Bubbles, Fads and Stock Price Volatility Tests: A Partial Evaluation," *Journal of Finance*, 43, 639–656.

Wu, G., 2001, "The Determinants of Asymmetric Volatility," *Review of Financial Studies*, 14, 837–859.