Aggregate Short Interest and Market Valuations

By Owen A. Lamont and Jeremy C. Stein*

The spectacular rise and fall of stock prices during the recent dot-com bubble period has been accompanied by a surge of interest in the topic of short-selling. For the most part, this work is cross-sectional in nature, examining the causes and consequences of short-sales constraints at the individual-stock level, and it suggests the following two broad conclusions. First, consistent with the notion that short-selling is undertaken by rational arbitrageurs, the demand for short positions is greatest among stocks that appear to be overvalued (e.g., stocks that have high ratios of prices to book value). Second, because of frictions in the market for borrowing stock, as well as various institutional rigidities, arbitrage by would-be short-sellers is incomplete. Thus, those stocks where the demand for shorting is greatest (as measured, say, by a high premium paid to borrow the stock for the purposes of short-selling) tend to have abnormally low future returns (see e.g., Patricia Dechow et al., 2001; Joseph Chen et al., 2002; Gene D'Avolio, 2002; Charles Jones and Owen Lamont, 2002; Lamont and Richard Thaler, 2003; Eli Ofek and Matthew Richardson, 2003).

Less attention has been paid to variation over time in aggregate short interest, and to the role that this might have in countering market-wide sentiment. Casual intuition might suggest that short-selling-based arbitrage would be *more* effective along the aggregate dimension than it is in the cross section. After all, while it can be difficult at any point in time to short a minority of very overpriced stocks, most stocks are easily and cheaply shorted. Moreover, there are other ways to get a short bet down on the aggregate market—for example, by purchasing put options on various indices.

It turns out that this intuition is off the mark. We examine some basic data on the evolution of aggregate short interest, both during the dot-com era, and at other times in history. In a striking contrast to the patterns seen in the cross section, total short interest moves in a countercyclical fashion. For example, short interest in NASDAQ stocks actually declines as the NASDAQ index approaches its peak. Moreover, this decline does not seem to reflect a substitution away from outright short-selling and toward put options: the ratio of put-to-call volume displays the same countercyclical tendency. As we discuss below, the evidence is perhaps most consistent with Andrei Shleifer and Robert Vishny (1997), who argue that the open-end nature of most professional arbitrage firms (i.e., the fact that investors can withdraw their funds on demand) makes it difficult for these firms to buck aggregate mispricings. The evidence also suggests that short-selling does not play a particularly helpful role in stabilizing the overall stock market.

I. The Data

A. *The Dot-Com Bubble*

Figure 1 tells our basic story for the dot-com period. We plot three series on a monthly basis over the interval 1995–2002: (i) the NASDAQ index (the Center for Research in Security Prices [CRSP]'s total return index); (ii) the value-weighted short-interest ratio (100 times the market value of shares sold short, divided by the value of shares outstanding) for all NASDAQ companies; and (iii) the 60-day moving average of the Chicago Board Options Exchange's (CBOE) daily put–call ratio. The put–call ratio is the total CBOE trading volume in puts (including both index options and options on individual NASDAQ, NYSE, and AMEX stocks) divided by the volume in calls, and we use it as an admittedly noisy proxy for the

* School of Management, Yale University, New Haven, CT 06520, and Department of Economics, Harvard University, Cambridge, MA 02138, respectively. Thanks to the National Science Foundation for financial support, to Stefan Nagel and John Griffin for providing some of the data, to Andrea Frazzini and James David for research assistance, and to Harrison Hong, Charles Jones and Andrei Shleifer for helpful comments.

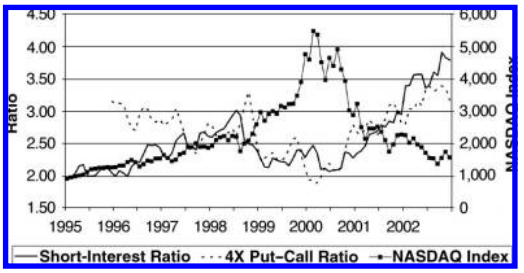

FIGURE 1. NASDAQ INDEX, NASDAQ SHORT-INTEREST RATIO, AND CBOE PUT–CALL RATIO, JANUARY 1995–DECEMBER 2002

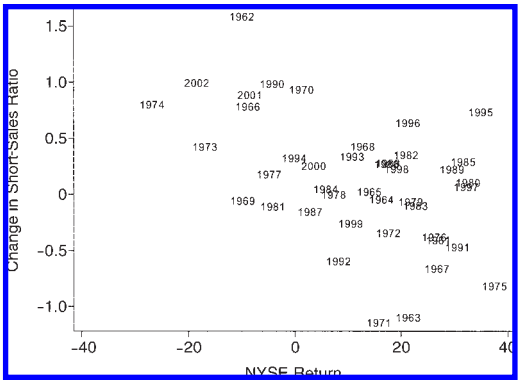

FIGURE 2. CHANGES IN NYSE SHORT-SALES RATIO VERSUS NYSE RETURN, 1961–2002

magnitude of shorting done via options. This ratio averaged about 0.7 during the period; we have multiplied it by 4 in the figure so as to fit it on the same scale as the short-interest ratio.

As can be seen, both the short-interest ratio and the put–call ratio decline substantially as the NASDAQ index explodes upward from mid-1998 to its peak in March 2000; they both then rebound sharply as the index collapses over the subsequent two years. Some simple statistics confirm the visual impressions from the figure. The return on the index over the prior 12 months has a correlation of −0.54 with the short-interest ratio; and a correlation of −0.63 with the put–call ratio. (The short-interest ratio and the put–call ratio are themselves highly positively correlated, at 0.58, suggesting that these two measures are capturing similar information.)

Aside from these time-series patterns, it is also worth noting that remarkably little short-selling takes place *at any point* in the cycle. In the case of the NASDAQ, the short-interest ratio averages 2.53 percent over our sample period, and it never breaks 4 percent.

B. *Short-Selling on the NYSE, 1960–2002*

For a longer historical perspective, we examine NYSE data from 1960 to 2002. Because of both data availability and institutional differences, we use an alternative measure of short-selling. One issue is that we have better data here on short-selling volume than open interest. A second is that on the NYSE, unlike the NASDAQ, a large fraction of shorting is due to specialists, who are engaged in high-frequency hedging. So the measure we use is total shares sold short by public investors (as opposed to by NYSE member firms) divided by total share volume, which we term the short-sales ratio, and which we can calculate on an annual basis.

The NYSE short-sales ratio trends sharply upward during this period (rising from 1.2 percent in 1960 to 6.6 percent in 2002), perhaps reflecting the growing popularity of hedge funds and other long–short strategies. Thus we look at the change in the short-sales ratio. In Figure 2, we plot this change against the annual return to the value-weighted NYSE stock index. The two series move strongly counter to one another: the correlation coefficient is −0.51, which is highly statistically significant. So overall, this longer stretch of history tells much the same story as Figure 1 does for the dot-com era.[1]

II. Implications

The evidence suggests that aggregate short interest displays *extrapolative* behavior; that is, it looks like fewer investors are willing to bet on the market going down after a period in which it has been rising. But this characterization

[1] Going further back in time, Jones and Lamont (2002) discuss the crash of 1929. Although quantitative data are scarce, anecdotal evidence indicates that, as stock prices rose in the late 1920's, short-selling declined. According to J. Edward Meeker (1932 p. 125), prior to the crash "few had the hardihood to sell short" and so "the panic of 1929 descended on an inadequate short interest."

raises a puzzle. Recall that at the individual-stock level, short interest appears to be *contrarian* in nature, with high-priced stocks attracting more attention from short-sellers. So why does the cross section of shorting seem to reflect the actions of rational arbitrageurs, while the aggregate time series seems to reflect the actions of naïve trend-chasers?

One potential answer has to do with the open-end nature of professional money management. Consider an example in the spirit of Andrei Shleifer and Robert Vishny (1997). Suppose there is a set of hedge funds that specialize in short-selling. The managers of these funds are rational arbitrageurs, so at any point in time, they will use the capital they have to target a portfolio of the most overvalued companies—hence, the pattern of shorting in the cross section. But when the market rises, the short-selling funds will lose money, and hence will face redemptions from their clients. These redemptions (i.e., the well-known "performance–flow" relationship) may have their roots in either rational updating about hedge-fund-manager ability, or in some degree of irrational trend-chasing on the part of end investors. But in either case, the result is that fund managers have less capital to work with in a rising market and are forced to scale back their aggregate short positions.

The broad message is that because of the pervasiveness of open-ending, professional arbitrage may be even less effective at countering market-wide sentiment shocks than it is at enforcing rational pricing in the cross section. This can be true even though the most obvious direct impediments to arbitrage (e.g., individual stocks being hard to borrow) arise in the cross section.[2]

Of course, this line of discussion raises another question: if open-ending is such a handicap for arbitrageurs when it comes to dealing with market-wide sentiment, why is the open-end form so common? On the one hand, it seems clear that open-ending is a natural response to problems of agency and asymmetric information. Investors worry about turning over their money to somebody who may turn out to be incompetent or crooked, and so they attach value to an early-liquidation option. Yet it does not follow that the degree of open-ending that we observe is one that serves investors well. Stein (2003) argues that competition among money managers for investors' dollars creates an externality and can lead to a socially excessive amount of open-ending. When any one fund open-ends, it compromises its own ability to undertake certain kinds of arbitrage (which is both a private and social cost), but it makes itself more attractive to investors, and thereby steals business from other funds (which is a private, but not a social gain).

This general perspective on the constraints faced by professional money managers is helpful in thinking about the arbitrage role played by nonfinancial firms. In contrast to the behavior documented above, nonfinancials were, effectively, enormous short-sellers during the bubble period, via issues of their own shares: the dollar volume of initial public offerings and seasoned equity offerings peaked at roughly the same time as the NASDAQ index. In rationalizing this fact, one probably does not want to take the position that nonfinancial managers are shrewder or better informed than, for example, hedge fund managers, particularly with respect to market-wide movements in prices. A more plausible explanation has to do with a comparative institutional advantage. A nonfinancial manager who issues equity at the time of a market peak does so in the closed-end corporate form, and without being subject to mark-to-market accounting. Thus, if the market continues to go up, she will not record a loss, and she will certainly not be faced with the threat of liquidation.

As a final point, our data shed some light on the tendency for short-sellers to come under political attack in the aftermath of large market declines. Jones and Lamont (2002) discuss the crackdown on short-selling after the crash of 1929 and note that numerous anti-shorting regulations stem from this period, including the uptick rule, as well as the Investment Company Act of 1940, which placed severe restrictions on the ability of mutual funds to go short. It is clear from Figures 1 and 2 that aggregate short-selling tends to increase in bear markets, which perhaps

[2] See also John Griffin et al. (2003) and Markus Brunnermeier and Stefan Nagel (2004). Both papers focus directly on the actions of large institutions and find that they actually had substantial long positions in high-priced tech stocks during the period in which the NASDAQ index was approaching its peak.

makes it all the easier for people to blame the messenger. However, according to our interpretation of this evidence, the problem is not too much short-selling in falling markets (recall that the aggregate volume of short interest is always quite small in absolute terms), but rather, too little in rising markets. If this view is correct, any regulatory efforts to constrain short-selling are likely to be misguided.

REFERENCES

Brunnermeier, Markus and Nagel, Stefan. "Hedge Funds and the Technology Bubble." *Journal of Finance*, 2004 (forthcoming).

Chen, Joseph; Hong, Harrison and Stein, Jeremy. "Breadth of Ownership and Stock Returns." *Journal of Financial Economics*, November 2002, *66*(2–3), pp. 171–205.

D'Avolio, Gene. "The Market for Borrowing Stock." *Journal of Financial Economics*, November 2002, *66*(2–3), pp. 271–306.

Dechow, Patricia; Hutton, Amy; Meulbroek, Lisa and Sloan, Richard. "Short-sellers, Fundamental Analysis, and Stock Returns." *Journal of Financial Economics*, July 2001, *61*(1), pp. 77–106.

Griffin, John; Harris, Jeffrey and Topaloglu, Selim. "Investor Behavior over the Rise and Fall of NASDAQ." Working paper, Yale University, 2003.

Jones, Charles and Lamont, Owen. "Short Sale Constraints and Stock Returns." *Journal of Financial Economics*, November 2002, *66*(2–3), pp. 207–39.

Lamont, Owen and Thaler, Richard. "Can the Market Add and Subtract? Mispricing in Tech Stock Carve-Outs." *Journal of Political Economy*, April 2003, *111*(2), pp. 227–68.

Meeker, J. Edward. *Short selling.* New York: Harper Brothers, 1932.

Ofek, Eli and Richardson, Matthew. "DotCom Mania: The Rise and Fall of Internet Stock Prices." *Journal of Finance*, June 2003, *58*(3), pp. 1113–38.

Shleifer, Andrei and Vishny, Robert. "The Limits of Arbitrage." *Journal of Finance*, March 1997, *52*(1), pp. 35–53.

Stein, Jeremy. "Why Are Most Funds Open-End? Competition and the Limits of Arbitrage." Working paper, Harvard University, 2003.