7-2005

The Information Content of Short Interest: A Natural Experiment

Tom Arnold
University of Richmond, tarnold@richmond.edu

Alexander W. Butler

Timothy Falcon Crack

Y. Zhang

Recommended Citation
Arnold, Tom, Alexander W. Butler, Timothy Falcon Crack, and Yan Zhang. "The Information Content of Short Interest: A Natural Experiment." *The Journal of Business* 78, no. 4 (July 2005): 1307-336. doi:10.1086/430861.

Tom Arnold
E. Claiborne Robins School of Business, University of Richmond

Alexander W. Butler
University of Texas at Dallas

Timothy Falcon Crack
University of Otago

Yan Zhang
State University of New York, Binghamton

The Information Content of Short Interest: A Natural Experiment*

An increase in the cost of selling short should increase the bearish information content of short interest announcements by driving relatively uninformed short sellers out of the market. We extend the Diamond and Verrecchia (1987) model to include short selling against the box and test the extended model using a natural experiment based around the Taxpayer Relief Act of 1997 (TRA97). TRA97 made short selling more costly for those shorting against the box. Consistent with the implications of our extended model, this increase in short-selling costs strengthens the negative relationship between short interest and subsequent stock price performance post TRA97.

I. Introduction

A short sale against the box is a transaction in which a trader holding a long position in an asset sells the same asset short but does not immediately deliver the long position to cover the short sale. In doing this, the trader neutralizes his or her exposure to fluctuations in the value of the asset, without liquidating the long position. Thus, unlike a regular short sale, which leaves the short seller with negative exposure to the stock, a short

* We thank Paul Asquith, Lee Ann Butler, Michael Highfield, Ken Hui, Robert Klemkosky, Robert Jennings, Shane Johnson, Adam Lei, Dixon Lin, Ed Miller, Chip Ryan, Laura Starks, James Weston, Karen Wruck, and seminar participants at Tulane University, Louisiana State University, University of Texas at San Antonio, Wake Forest University, University of New Orleans, and College of William and Mary's Young Scholars Conference for helpful comments and suggestions.We thank Carol Bitting and Susan Farkas of the NYSE for providing data.Special thanks go to Todd Pulvino and Mark Mitchell for supplying merger data. Any errors are ours. Contact the corresponding author, Alexander Butler at butler@utdallas.edu.

(*Journal of Business,* 2005, vol. 78, no. 4)

0021-9398/2005/7804-0006$10.00

sale against the box leaves the trader with zero net exposure to the stock. Short selling is more expensive than selling outright due to various restrictions on short sales, including the uptick rule, margin requirements, payment of dividends, and restricted use of proceeds (see Alexander and Peterson 1999 and D'Avolio 2002). Why would a trader engage in this strategy rather than simply selling the asset outright to achieve the same nonexposure? One important reason is that, prior to the Taxpayer Relief Act of 1997 (TRA97), a short sale against the box allowed investors to eliminate their exposure to an appreciated financial position and preserve a capital gain while postponing the capital gains tax until a later tax year.[1]

TRA97 eliminated this opportunity to defer capital gains taxes by making a short sale against the box a "constructive sale" upon which capital gains or losses are immediately recognized unless some strict rules are met (Tucker and Watson 1999).[2] We formally extend the Diamond and Verrecchia (1987) model of the informativeness of short interest to allow for short selling against the box and use this change in the tax law as a natural experiment to test the prediction of our model extension.

Recent empirical work suggests that short interest is a bearish indicator, conveying negative sentiment (Figlewski and Webb 1993; Senchack and Starks 1993; Asquith and Meulbroek 1995; Aitken et al. 1998; Desai et al. 2002). Short selling contains negative information because short selling is relatively costly and subject to more restrictions than selling shares outright; consequently, short selling is more likely to be done by informed traders (Diamond and Verrecchia 1987).[3] Short selling is not limited to informed traders, however, and to the extent that uninformed or "noise" traders use short selling, we expect short interest in a stock to be a noisy signal of bearish market sentiment for that stock. The Diamond and Verrecchia (1987) model demonstrates that an increase in the cost of short selling increases the negative information contained in short interest announcements by driving relatively uninformed short sellers out

1. Before 1950, an investor could short sell against the box and lock in a short-term capital gain that would then be taxed at the lower long-term capital gains rate if the short position was held long enough. The ability to convert short-term capital gains to long-term ones was removed by the Revenue Act of 1950 (Macaulay and Durand 1951; Brent, Morse, and Stice 1990). However, until 1997, a trader who held a stock that had appreciated could lock in the gain and immunize price risk by selling short against the box. Thus, taxes could be delayed by holding the short position until the following tax year, at which time the long shares could be given to the stock lender to close the short position.

2. TRA97 made several changes to the tax code, including changing the maximum capital gains tax rate from 28% to 20% on long-term gains, excluding from taxes the first $500,000 in gain from the sale of a principal residence, and exempting from taxes the gains from the sale of small business stock (less than and $50 million in gross assets) if the proceeds are used to purchase other small business stock. We restrict our attention to the changes requiring investors to recognize gains from short selling against the box.

3. For institutional details of short selling, see Asquith and Meulbroek (1995), Angel (1997), Dechow et al. (2001), and D'Avolio (2002).

of the market at a higher rate than relatively informed short sellers. TRA97 provides an exogenous change in short selling costs, which we use to test this theory.[4]

Using a large sample of short interest announcements, we document the determinants of short interest and demonstrate that, prior to TRA97, short selling against the box was a popular trading strategy. TRA97 eliminated the tax benefits of short selling against the box, thereby making this trading strategy more costly. Diamond and Verrecchia (1987) predict that an increase in the costs of short selling should increase the information content of short interest announcements; our findings provide strong evidence in support of this view. One of the strengths of our paper over previous studies of the informativeness of short-interest announcements lies in the generality of our results. By exploiting an exogenous event as a natural experiment, we are able to demonstrate that the costs of short selling increase the negative information content of short interest announcements in general, rather than for just a subset of stocks, such as those that have very high levels of short interest (Asquith and Meulbroek 1995; Desai et al. 2002), traded options (Senchack and Starks 1993), or have an introduction of traded options (Danielson and Sorescu 2001).

Particularly for NYSE stocks, short-interest announcements after TRA97 convey significantly more negative sentiment than those before TRA97. In contrast to the findings of Asquith and Meulbroek (1995) and Desai et al. (2002), it is no longer true that the negative relationship between short interest and subsequent stock price performance holds only in the most heavily shorted stocks. Rather, we find that, after TRA97, short-interest announcements convey negative information even for stocks with moderate levels of short interest.

Announcements of high levels of short interest convey negative information for NASDAQ stocks both before and after TRA97. In contrast to the predictions of Diamond and Verrecchia, however, and in direct contrast to our findings for NYSE stocks, the highest levels of short interest in NASDAQ stocks have less negative information content after TRA97. One possible reason for this stems from an important institutional difference between the two markets—reforms to the NASDAQ market enacted during 1997 substantially decreased trading costs for

4. Because short selling is more costly than selling outright, we posit that short selling against the box is primarily a tax-motivated strategy rather than an information-motivated strategy. It is possible, however, that informed traders could use short selling against the box to hide information-based trades among those of the uninformed or that informed traders have both tax- and information-motivated reasons to eliminate downside risk in a stock. We postulate that the proportion of informed traders using short sales against the box is substantially smaller than the proportion of informed traders in the general short-selling population. After TRA97, the extra tax costs of short selling against the box make it more likely that, when a short sale occurs, it is a result of relatively informed traders with negative expectations, rather than due to tax motivations. We model this formally in Section III as an extension of Diamond and Verrecchia (1987).

large and mid-cap NASDAQ stocks (see Weston 2000). We speculate that this reduction in trading costs brought in uninformed traders in our high short-interest stocks, thereby muting the noise-reducing effect of TRA97.

We use analytical results from Dyl (1978) to establish the prevalence of short selling against the box prior to TRA97, then we provide a direct test of our extension of Diamond and Verrecchia (1987). Our paper is most closely related to the empirical work of Asquith and Meulbroek (1995) and Senchack and Starks (1993). Asquith and Meulbroek (1995) and Desai et al. (2002) demonstrate the negative relationship between high levels of short interest and subsequent stock returns. We demonstrate similar results, although we find that, following TRA97, the negative relation extends well beyond the highest short-interest stocks and to stocks with and without options.

The remainder of our paper is organized as follows. Section II describes literature on the costs and determinants of short selling. Section III presents our extension to the Diamond and Verrecchia (1987) model. Section IV discusses our empirical hypothesis. Section V describes our data, empirical methods, and results. Section VI presents our main result—that short interest is more informative after TRA97. Section VII concludes.

II. Costs and Determinants of Short Selling

The typical stock has very little short interest; most stocks have less than 0.5% of their shares outstanding held short. Thus, while there is substantial cross-sectional variation in short-interest levels, based in part on the determinants discussed in this section, the reader should bear in mind that short selling represents only a small proportion of total transactions in the average stock.

The pre-TRA97 determinants of tax-motivated short sales are described in Dyl (1978).[5] An investor who wishes to realize a capital gain must choose between selling the position outright and selling short against the box. By selling outright, the investor receives the proceeds from the sale immediately but incurs capital gains taxes in the current year. By selling short against the box, the investor defers the receipt of the proceeds because they are held as collateral against the short position by the broker. The investor also defers the capital gains taxes until a subsequent tax year. Thus, the after-tax net present value from selling outright is, using Dyl's notation, $P - T(P - C)$, where P is the current price of the stock, T is the applicable tax rate, and C is the cost basis for the investor's holdings. The after-tax net present value of selling short against the box is $P/(1+i)^{y} - T(P - C)/(1+i)^{12}$, where i is the monthly

5. This paragraph draws heavily upon Dyl (1978). We simplify the presentation of his model by assuming tax rates do not change.

opportunity cost of funds and y is the number of months after the short sale that the short position is covered. The relevant opportunity cost of funds is the monthly cost of borrowing the stock to short (see Reed 2002 for a comprehensive discussion of stock lending rates). The proceeds (P) are discounted until the short position is covered; the tax liability is discounted a full 12 months, because deferring the recognition of the tax defers payment for a full year: for instance, taxes on capital gains in 1995 are paid on or about April 15, 1996; deferring recognition of those gains to the 1996 tax year means they are paid on or about April 15, 1997.[6] The tax-savvy investor thus trades off the two net present values described previously when deciding whether to sell outright or sell short against the box. Key comparative statics results that obtain are that lower tax rates (T), longer required holding periods for the short position (y), and smaller capital gains ($P - C$) all reduce the relative profitability of selling short against the box, ceteris paribus. We draw upon these two latter results in Sections IV and V. In those sections, we argue, first, that a minimum cost strategy is likely to involve shorting as late as possible in the tax year and covering as soon as possible in the new tax year and, second, that the higher is the cumulative stock return, the more likely is a short sale.

Jones and Lamont (2002) argue that using short interest as a proxy for shorting demand is problematic because the costs of short selling, rather than the realized amount of short selling, should be related to future stock performance. For example, in a stock where the cost of shorting is prohibitively high, there is likely to be little or no short interest, yet stock prices are likely to be optimistic (as in Miller 1977) and subsequent stock performance is likely to be poor. Thus, to properly test the relation between short interest and future returns, it is important to look at the marginal impact of changes in the cost of short selling on future returns; this is exactly what we do.[7]

To attribute changes in short-selling behavior to the changes in short-selling costs brought about by TRA97, however, we need to control for non-information-based determinants of short selling. Diamond and Verrecchia (1987) argue that introducing option contracts on a stock reduces the costs associated with short selling, because option strategies

6. To put the effect of TRA97 into the context of the model, the tax liabilities are now discounted for 0 months instead of 12 (because a short sale against the box is treated as a "constructive sale"), which means that there are no net benefits to selling short against the box after TRA97.

7. Although the use of short interest as a determinant of future stock return cannot escape this criticism, short interest data are much more readily available than explicit shorting costs, however; and ignoring the few hot stocks that are prohibitively costly to short, our use of short interest is justified. Indeed, D'Avolio (2002, p. 273) notes that 91% of stocks in his U.S. sample are cheap to borrow (the so-called general collateral stocks), 8% of his sample are lent out at more expensive "special" rates, and the final 1% comprise hot stocks with very high loan rates.

allow traders to mimic short-selling strategies.[8] Figlewski and Webb (1993) and Senchack and Starks (1993) present empirical evidence that optionable stocks display more informational efficiency than nonoptionable stocks. Brent et al. (1990), Aitken et al. (1998), and Graham, Hughen, and McDonald (1999) all use dummy variables to control for whether a stock is optionable or not. They argue that the use of options is associated with the use of short selling for hedging and options arbitrage purposes, which must be controlled for when looking at information-based reasons for short selling.

Graham et al. (1999) include the bid-ask spread as a proxy for trading costs, asserting that traders are more likely to short sell stocks for which the trading costs are lower. We use average trading volume as an alternate proxy for trading costs—higher volume stocks are more liquid, thus it is less costly to open and close short positions. We do not expect bid-ask spreads to systematically affect our results. In fact, due to reductions in minimum price variation, effective bid-ask spreads have decreased on average over our sample period (see Goldstein and Kavajecz 2000 and Bessembinder 2000), biasing against our finding that the informativeness of short-interest announcements has increased post TRA97.

Dechow et al. (2001) argue that dividend yields and institutional ownership are proxies for trading costs, because dividends must be paid by the investor who holds the short position and higher institutional ownership reflects a larger pool of shares that can be borrowed. They also control for market capitalization, because it is easier to borrow a large-capitalization stock. Average dividend yields have decreased during our sample period, making short selling less costly in general. We control for such systematic determinants of short interest by including aggregate short interest in our regressions (see later).

According to Sias and Starks (1997), capitalization and institutional ownership are highly correlated. To the extent that large capitalization companies have a large number of shares held by institutional investors, we can indirectly control for the institutional ownership by including market capitalization in our regressions (see later). Dechow et al. (2001) state that short sellers target firms with prices that are temporarily high relative to fundamentals and in which it is expected that prices, not fundamentals, will revert. According to our hypotheses, trades enacted under these trading strategies have more impact after TRA97.

Brent et al. (1990) and Graham et al. (1999) also include dummies for whether a company has convertible debt or not. Because convertible debt has an imbedded warrant on the underlying stock, convertible bond

8. Note, however, that Asquith and Meulbroek (1995) state that discussions with short sellers reveal that they do not consider options trades as an alternative to short selling because the costs of option trading are too high (p. 6). Miller (1977, p. 1161) makes a similar point regarding the high costs of option trading.

holders (or bond arbitrageurs) may wish to hold short positions in the stock to offset the imbedded warrant. Brent et al. (1990) also argue that investors are more likely to short against the box if the stock is volatile. However, they find no statistical evidence of a relationship between the stock beta and short interest or between the market-model residual variance and short interest.

Recent studies document that short selling is also related to merger activity (Mitchell and Pulvino 2001; Cornelli and Li 2002; Mitchell, Pulvino, and Stafford 2003). Risk arbitrageurs attempt to profit from the spread between the target stock price and the offer price. In a stock merger, arbitrageurs lock in the spread by taking a long position in the target stock and selling short the acquirer's stock. Indeed, Mitchell et al. (2003) claim that almost half of the fall in stock price experienced by acquirers during stock-financed merger announcement periods is due to price pressure from short sellers engaged in merger arbitrage.

Of the above-mentioned determinants of short interest, we directly control for volume, options, market value, convertible securities, and merger arbitrage. We also control for other variables not yet mentioned, as explained in Section V. Our findings are consistent with prior studies: the existence of options on a stock or convertible securities is a major determinant of short interest. Similarly, trading volume and merger activity are strongly related to short-selling activity.

III. Extension of Diamond and Verrecchia (1987)

The Diamond and Verrecchia model excludes the possibility that investors who hold shares long may want to sell short. In their model, this is reasonable because selling short is costly relative to selling outright, and there are no tax motivations to trade in their model. We offer a simple extension of their model, where some investors may want to sell short against the box, that is, sell short an asset they also hold long, as opposed to selling the asset outright. We first list our notation in tables 1 and 2.

To facilitate comparison to the Diamond and Verrecchia model, we use their notation and model setup throughout. Figure 1 replicates the game tree from Diamond and Verrecchia (their figure 1). Figure 2 adds our extension to the Diamond and Verrecchia model. The salient difference between our extension and the Diamond and Verrecchia model is that a trader who holds shares long might sell the shares or short sell the shares (while maintaining the long position); we explicitly allow short selling against the box in our extension, but Diamond and Verrecchia abstract from tax-motivated reasons to short sell. We denote the probability of a tax benefit to short selling against the box (conditional upon the trader wanting to sell or short sell and holding shares long) as q.

As in Diamond and Verrecchia, we allow three costs of short selling: no cost (or, without loss of generality, minimal or trivial cost), restrictive

TABLE 1 Notation

Variable	Definition
v	Value of the asset, either 1 or 0.
g	Probability that one trader wants to trade (for liquidity- or information-based motives).
a	Probability that a given trader is informed (the fraction of traders who are informed).
h	Probability that a trader already owns the stock (the fraction of traders who hold long positions).
C_i	Probability that a trader faces cost i of short selling (the fraction of traders facing cost i). C_1 denotes no costs to sell short, C_2 denotes modest costs to sell short, C_3 denotes prohibitive costs to sell short; $C_1 + C_2 + C_3 = 1$.
q	Probability that a trader who holds shares will have tax benefits from selling short against the box.

costs of short selling (modeled by Diamond and Verrecchia as deferred receipt of proceeds, but without loss of generality, this can be any nonprohibitive cost of short selling), and prohibitive costs. Following Diamond and Verrecchia, we denote the fraction of traders facing these costs as C_1, C_2, and C_3, respectively, and $C_1 + C_2 + C_3 = 1$. These costs are such that informed traders with bad news find it profitable to sell short if they are of types 1 or 2 but not of type 3; uninformed traders with liquidity motivated reasons to sell will sell short only if they are type-1 traders.

When there are tax-deferring advantages to short selling against the box, some traders may prefer selling short against the box to selling outright if the tax deferring benefit exceeds the costs of selling short. We assume both informed and uninformed traders face the same tax benefits and denote the proportion of traders with a tax benefit to selling short against the box as q and those choosing to sell outright as $1 - q$ (both conditional upon holding shares). Now, the tax benefits from a short sale against the box can exceed the costs of the short sale for some traders (i.e., those C_1 and C_2 traders in group q), inducing some traders to sell short against the box because of the tax effects. If the benefits from a short sale against the box do not exceed the costs of the short sale (i.e., for those C_3 traders in group q), traders sell outright rather than sell short against the

TABLE 2 Types of Traders and How They Trade

Cost Type	Informed with Bad News and No Long Position	Uninformed with Preference to Sell and No Long Position	Informed with Bad News and Holds Shares	Uninformed with Preference to Sell and Holds Shares
C_1: No cost	Short	Short	Sell or SSAB	Sell or SSAB
C_2: Restrictive costs	Short	No trade	Sell or SSAB	Sell or SSAB
C_3: Prohibitive costs	No trade	No trade	Sell	Sell

NOTE.—Short selling against the box is denoted SSAB.

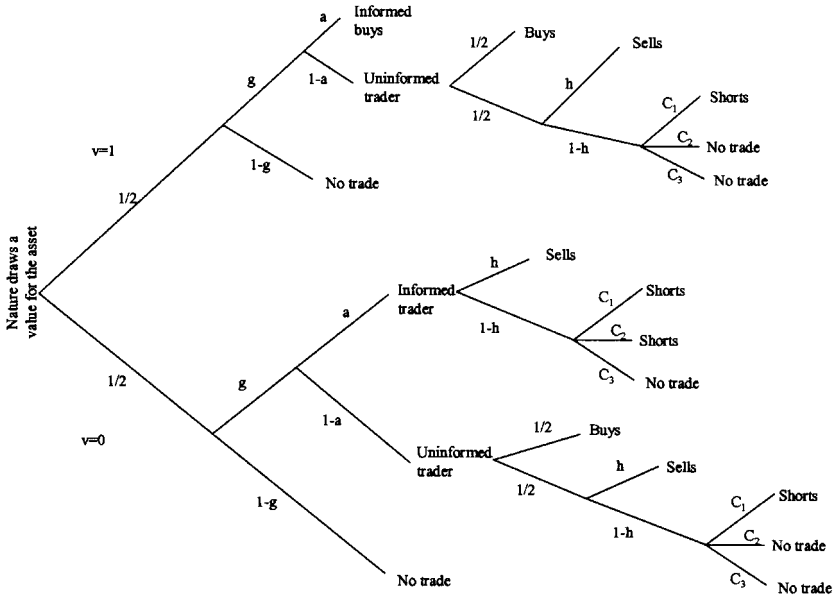

FIG. 1.—The Diamond and Verrecchia (1987) model. This figure replicates the tree diagram of the model from Diamond and Verrecchia (1987, fig. 1) and so does not allow for short selling against the box.

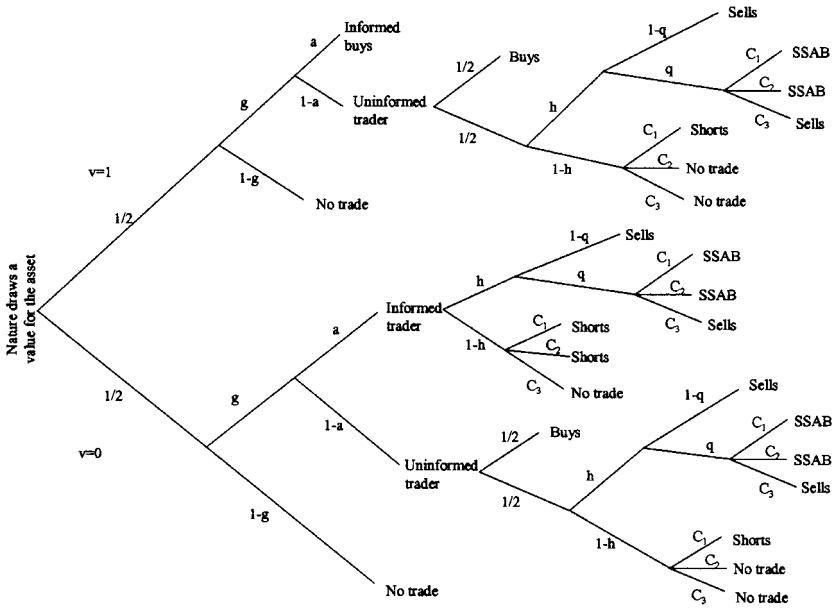

FIG. 2.—Extension of Diamond and Verrecchia (1987). This figure presents the tree diagram of the extended version of the model from Diamond and Verrecchia (1987), where short selling against the box (denoted SSAB in the figure) is possible.

box (see the summary in table 2). Note that, if q is zero, indicating that no investors enjoy a tax benefit to selling short against the box, our extension collapses to the original Diamond and Verrecchia model (figure 1).

We use this simple extension to examine the effects of eliminating short selling against the box, such as with the Taxpayer Relief Act of 1997. Specifically, we derive two key theoretical results (derived and proved in the appendix).

Proposition 1 (Reduction of Short Selling against the Box). As the probability of tax benefits to short selling against the box, q, decreases, the proportion of informed short selling increases.

Corollary to Proposition 1 (Elimination of Short Selling against the Box). If q is zero, the proportion of informed short selling is greater than if q were strictly positive.

Taken together, these two results provide an important empirical prediction: if markets or regulations change in such a way that the proportion of traders wanting to sell short against the box decreases, the proportion of short sellers who are informed traders increases. This, in turn, increases the negative information content in short-interest announcements. We test this prediction in Section VI.

IV. Hypothesis

Our hypothesis is that the information content of short interest increases after TRA97 due to the departure from the market of (relatively uninformed) tax-motivated traders selling short against the box. This hypothesis stems directly from our extension of Diamond and Verrecchia (1987). Before testing our hypothesis, we first confirm that there was a change in short selling against the box after the implementation of TRA97. Having done that, we then compare the information content of short interest before and after TRA97.

Our data identify short interest in individual stocks, but short sales against the box are not directly identifiable as such. We identify short selling against the box indirectly by postulating a "minimum-cost strategy" that pre-TRA97 tax-motivated traders selling short against the box would likely use, based on Dyl (1978). The logical strategy to defer taxes and minimize trading costs prior to TRA97 is to sell an appreciated stock short against the box before or during December and close the position in the following January. A taxpayer following this strategy could have locked in gains on the appreciated stock, postponed paying capital gains taxes until the following tax season, and maintained the costly short position for as short a time as possible—a minimum-cost strategy.[9] Traders

9. This is precisely the strategy advocated by Dyl's (1978) deterministic model of when an investor should sell short against the box instead of sell the stock outright—investors want to minimize y, the length of time the short position is open, ceteris paribus.

could implement or close short sales against the box positions at other times of the year.[10] Short selling is costly, however, so we expect rational investors to maintain short positions for the shortest time period possible, ceteris paribus. If taxpayers followed this minimum-cost strategy prior to TRA97, then we should see abnormally large short interest in December as investors implement their tax-deferring strategies and substantially smaller short interest in January as investors close their tax-deferring positions. Examination of how this phenomenon changed after the implementation of TRA97 gives a conservative estimate of the effect of TRA97 upon levels of short selling against the box.

In Section V, we identify trading patterns consistent with short selling against the box. In Section VI, we test our main hypothesis, that the negative information contained in short interest announcements increased after TRA97.

V. Data, Methods, and Results

A. Sample Description

Short-interest data from January 1995 through November 1999 come from the NYSE and from the NASDAQ Web site.[11] NYSE and NASDAQ short-interest data are compiled monthly, based on short positions held as of settlement on the 15th of each month; the data are made publicly available several days later (4 business days later for NYSE, 8 business days later for NASDAQ). Because we examine short interest in December and January to infer levels of short selling against the box, one potential concern is that any short positions implemented after the 15th of a given month and closed prior to the 15th of the following month do not appear in our data; this "invisible" short selling biases against our finding significant results.

In a typical month, well over 5,000 NASDAQ and 3,000 NYSE stocks and other securities are reported with short interest.[12] Most prior studies of short interest examine samples of 200 or fewer stocks or they examine stocks with unusually high levels of short interest. The latter increases the power of their tests but allows researchers to draw few conclusions about stocks with more typical levels of short interest. We examine a much larger sample, allowing us to make broader, more general inferences.

We include in our study those common stocks that have reported short interest (including zero short interest) for each month during our

10. If a trader opened a short sale against the box position prior to December to defer realization of capital gains taxes, he or she would need to hold the position through the end of the tax period and so would still hold the short position in December, regardless of when the short position was opened.

11. See www.marketdata.NASDAQ.com for NASDAQ data details.

12. These include common stock, preferred stock, and warrants. Approximately 75% are common stock.

sample period, resulting in a sample of 1,129 NASDAQ stocks and 535 NYSE stocks for which we have price and return information from the Center for Research in Security Prices (CRSP).[13] This increases the power of our tests without being as restrictive as Asquith and Meulbroek (1995), for example, who primarily examine only heavily shorted stocks. To mitigate any confounding effects of TRA97 being implemented midmonth (June 8, 1997), we exclude June 1997 observations from our sample. Descriptive statistics for our sample are in tables 3, 4, and 5.

Table 3 describes for our NYSE data the distribution of aggregate short interest for our sample, short interest as a percentage of total shares outstanding (i.e., relative short interest), index and security returns, trading volume, the presence of options and convertible securities, market value of equity for our sample stocks, and a dummy for whether the firm is in the process of acquiring another. We report descriptive statistics for these variables for the full sample and stratify by how much relative short interest each stock has. We segregate our sample into quintiles of relative short interest, with Quintile 1 having the least relative short interest, and Quintile 5 the highest relative short interest. Table 4 presents analogous information for our NASDAQ data.

Table 5 presents correlations among our variables. Our measures of short interest are strongly correlated with firm size, trading volume, the presence of options or convertible securities, and merger activity.

Figure 3 shows time trends of short interest for the top three relative short-interest quintiles for each exchange. The average relative short interest of each NASDAQ quintile is higher than the average relative short interest of the corresponding NYSE quintile for every month in our sample. The graph shows a fairly consistently upward trend in short interest from 1995 through 1998; relative short interest drops during 1999 for each quintile of short interest. The peak corresponds roughly with the month immediately prior to the collapse of the infamous hedge fund Long-Term Capital Management (in late September 1998).

B. Determinants of Short Interest and Short Selling against the Box

We perform two sets of regressions. In the first set of regressions, we identify the determinants of short interest over the sample period and assess the extent of short selling against the box. In the second set of regressions, we examine the information content of short interest over the sample period (see Section VI). By examining the differences between the regression coefficients in the regressions, we can determine the effects of TRA97.

13. We exclude all those securities that are not common stocks, which eliminates about 25% of the original sample, and those are not listed in CRSP, which eliminates about 30% of the original sample. Our requirement of a full, balanced panel of data is the most restrictive of our screens and brings us to our final sample size of 1,664 stocks.

TABLE 3 NYSE Descriptive Statistics

		Size	Volume	Stock Return	Market Return	Option	Relative Short Interest	Log Natural (Size)	Convertibles	Risk Arbitrage	Aggregate Short Interest
Full sample	Mean	3,960,113	256,563	1.16%	1.98%	.4935	.013425	13.2637	.1300	.0286	659.72
	Median	431,818	50,552	.81%	2.43%	.0000	.003980	12.9758	.0000	.0000	718.85
	SD	13,765,537	656,417	10.20%	4.55%	.5000	.030314	1.8305	.3363	.1667	183.68
Least relative short	Mean	397,357	33,683	1.29%	1.98%	.1018	.000136	12.0240	.0649	.0069	
interest quintile	Median	152,430	15,402	.77%	2.43%	.0000	.000089	11.9345	.0000	.0000	
(Quintile 1)	SD	1,287,848	61,042	9.17%	4.55%	.3025	.000143	1.1224	.2464	.0830	
Quintile 2	Mean	767,477	53,028	.86%	1.98%	.2054	.001060	12.3901	.0799	.0056	
	Median	213,023	22,759	.67%	2.43%	.0000	.000939	12.2692	.0000	.0000	
	SD	3,756,971	165,631	9.00%	4.55%	.4041	.000574	1.1983	.2712	.0749	
Quintile 3	Mean	6,702,354	304,139	1.12%	1.98%	.5390	.004166	13.5883	.1076	.0150	
	Median	576,690	53,406	.82%	2.43%	1.0000	.003980	13.2651	.0000	.0000	
	SD	19,536,904	810,761	9.66%	4.55%	.4985	.001683	1.9884	.3099	.1215	
Quintile 4	Mean	8,113,917	447,355	1.47%	1.98%	.7955	.011157	14.3866	.1481	.0358	
	Median	1,726,825	165,917	1.24%	2.43%	1.0000	.010495	14.3618	.0000	.0000	
	SD	19,979,081	851,674	10.40%	4.55%	.4034	.003680	1.8315	.3552	.1857	
Most relative short	Mean	3,819,457	444,610	1.08%	1.98%	.8255	.050604	13.9296	.2493	.0798	
interest quintile	Median	1,168,485	206,876	.66%	2.43%	1.0000	.033009	13.9712	.0000	.0000	
(Quintile 5)	SD	10,161,538	758,748	12.39%	4.55%	.3796	.052684	1.6250	.4326	.2709	

NOTE.—This table presents descriptive statistics of sample data. Data are for the full sample of 535 NYSE stocks during the period January 1995 through November 1999. Descriptive statistics are reported for firm size (in thousands of dollars), volume (the average daily volume over the previous month), monthly stock return, CRSP value-weighted index return, dummy variable for options (1 = traded options on the stock, 0 otherwise), relative short interest (total number of shares held short divided by total number of shares outstanding), natural log of firm size, dummy variable for convertibles (1 = convertible securities in the firm's capital structure, 0 otherwise), dummy variable for risk arbitrage (1 = firm has announced an acquisition of another publicly traded firm; changes to 0 when the target delists; 0 otherwise), and the aggregate short interest for our sample firms in a given month (in millions of shares).

TABLE 4 NASDAQ Descriptive Statistics

		Size	Volume	Stock Return	Market Return	Option	Relative Short Interest	Log Natural (Size)	Convertibles	Risk Arbitrage	Aggregate Short Interest
Full sample	Mean	900,439	242,424	2.15%	1.98%	.4460	.017029	11.9079	.1157	.0132	629.39
	Median	133,400	50,946	.04%	2.43%	.0000	.004730	11.8011	.0000	.0000	665.10
	SD	9,200,211	1,061,939	18.56%	4.55%	.4971	.032383	1.5087	.3198	.1142	218.43
Least relative short	Mean	130,269	30,435	2.49%	1.98%	.1211	.000228	11.1099	.0925	.0050	
interest quintile	Median	66,884	14,729	.59%	2.43%	.0000	.000185	11.1107	.0000	.0000	
(Quintile 1)	SD	214,153	66,010	16.54%	4.55%	.3263	.000192	1.1522	.2897	.0708	
Quintile 2	Mean	219,384	57,595	2.46%	1.98%	.2731	.001489	11.3765	.0797	.0051	
	Median	85,399	29,152	.06%	2.43%	.0000	.001334	11.3551	.0000	.0000	
	SD	3,799,257	221,370	18.39%	4.55%	.4456	.000729	1.2206	.2709	.0713	
Quintile 3	Mean	935,746	143,810	2.05%	1.98%	.4513	.005076	11.8893	.0935	.0095	
	Median	133,924	52,021	.09%	2.43%	.0000	.004751	11.8050	.0000	.0000	
	SD	14,080,371	923,293	17.96%	4.55%	.4976	.001995	1.3895	.2912	.0972	
Quintile 4	Mean	2,220,055	454,313	1.92%	1.98%	.6322	.014508	12.4637	.1015	.0209	
	Median	218,577	103,150	.03%	2.43%	1.0000	.013541	12.2949	.0000	.0000	
	SD	14,036,965	1,796,285	18.90%	4.55%	.4822	.005326	1.6481	.3021	.1429	
Most relative short	Mean	996,740	525,967	1.82%	1.98%	.7524	.063844	12.7001	.2112	.0256	
interest quintile	Median	312,512	210,637	.00%	2.43%	1.0000	.045961	12.6524	.0000	.0000	
(Quintile 5)	SD	3,265,377	1,136,799	20.75%	4.55%	.4317	.048435	1.4356	.4082	.1580	

NOTE.—This table presents descriptive statistics of sample data. Data are for the full sample of 1,129 NASDAQ stocks during the period January 1995 through November 1999. Descriptive statistics are reported for firm size (in thousands of dollars), volume (the average daily volume over the previous month), monthly stock return, CRSP value-weighted index return, dummy variable for options (1 = traded options on the stock, 0 otherwise), relative short interest (total number of shares held short divided by total number of shares outstanding), natural log of firm size, dummy variable for convertibles (1 = convertible securities in the firm's capital structure, 0 otherwise), dummy variable for risk arbitrage (1 = firm has announced an acquisition of another publicly traded firm; changes to 0 when the target delists; 0 otherwise), and the aggregate short interest for our sample firms in a given month (in millions of shares).

TABLE 5 Correlations among Variables

	Size	Natural Log (Size)	Volume	Natural Log (Volume)	Short Interest	Natural Log (Short Interest)	Relative Short Interest	Agg. Short Interest, Both Exchanges	Stock Return	Market Return	Option	Convertibles	Risk Arbitrage
Size	1.0000	.4410	.6820	.3206	.6143	.2400	−.0003	.0624	.0146	−.0003	.1696	.0013	.1254
		.0001	.0001	.0001	.0001	.0001	.9321	.0001	.0001	.9323	.0001	.6792	.0001
Natural log (size)	.4401	1.0000	.4143	.6385	.4778	.6052	.1647	.0838	.0464	−.0027	.5889	.0124	.1552
	.0001		.0001	.0001	.0001	.0001	.0001	.0001	.0001	.3974	.0001	.0001	.0001
Volume	.6820	.4143	1.0000	.4730	.6732	.3154	.1280	.0758	.0276	.0003	.2256	.0315	.1067
	.0001	.0001		.0001	.0001	.0001	.0001	.0001	.0001	.9196	.0001	.0001	.0001
Natural log (volume)	.3206	.6385	.4730	1.0000	.4724	.7230	.3580	.0968	.0795	.0099	.6204	.1041	.1142
	.0001	.0001	.0001		.0001	.0001	.0001	.0001	.0001	.0022	.0001	.0001	.0001
Short interest	.6143	.4778	.6732	.4724	1.0000	.4264	.3510	.0903	.0069	−.0081	.2793	.0767	.1964
	.0001	.0001	.0001	.0001		.0001	.0001	.0001	.0315	.0118	.0001	.0001	.0001
Natural log (Short Interest)	.2400	.6052	.3154	.7230	.4264	1.0000	.4959	.1150	.0289	−.0163	.5747	.1236	.1287
	.0001	.0001	.0001	.0001	.0001		.0001	.0001	.0001	.0001	.0001	.0001	.0001
Relative short interest	−.0003	.1647	.1280	.3580	.3510	.4959	1.0000	.0493	.0068	−.0075	.2878	.1747	.0743
	.9321	.0001	.0001	.0001	.0001	.0001		.0001	.0346	.0205	.0001	.0001	.0001
Agg. short interest, both exchanges	.0624	.0838	.0758	.0968	.0903	.1150	.0493	1.0000	−.0567	−.0872	.0000	−.0598	.0344
	.0001	.0001	.0001	.0001	.0001	.0001	.0001		.0001	.0001	1.0000	.0001	.0001
Stock return	.0146	.0464	.0276	.0795	.0069	.0289	.0068	−.0567	1.0000	.2989	.0305	−.0035	−.0082
	.0001	.0001	.0001	.0001	.0315	.0001	.0346	.0001		.0001	.0001	.2818	.0108
Market return	−.0003	−.0027	.0003	.0099	−.0081	−.0163	−.0075	−.0872	.2989	1.0000	−.0001	−.0052	−.0114
	.9323	.3974	.9196	.0022	.0118	.0001	.0205	.0001	.0001		.9781	.1036	.0004
Option	.1696	.5889	.2256	.6204	.2793	.5747	.2878	.0000	.0305	−.0001	1.0000	.0855	.0779
	.0001	.0001	.0001	.0001	.0001	.0001	.0001	1.0000	.0001	.9781		.0001	.0001
Convertibles	.0013	.0124	.0315	.1041	.0767	.1236	.1747	−.0598	−.0035	−.0052	.0855	1.0000	.0159
	.6792	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.2818	.1036	.0001		.0001
Risk arbitrage	.1254	.1552	.1067	.1142	.1964	.1287	.0743	.0344	−.0082	−.0114	.0779	.0159	1.0000
	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.1008	.0004	.0001	.0001	

NOTE.—This table reports correlations among variables of interest and the *p*-value for a two-sided hypothesis test that each correlation is different from zero (reported below each correlation). Data are monthly observations for 535 NYSE stocks and 1,129 NASDAQ stocks during the period January 1995 through November 1999.

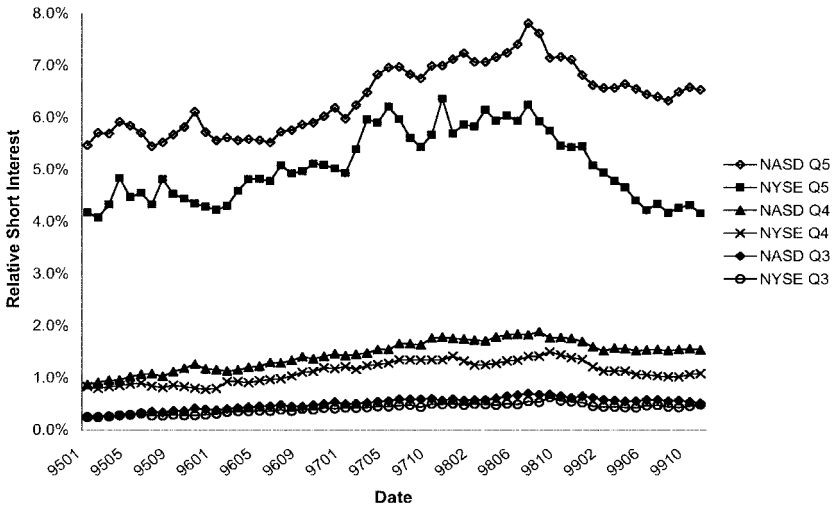

FIG. 3.—Evolution of aggregate NYSE and NASDAQ short interest. This figure shows the time series of relative short interest for our sample. Shown is the average short interest divided by shares outstanding (relative short interest) for the top three quintiles of relative short interest for both the NYSE and NASDAQ stocks in our sample.

We examine the determinants of short selling and whether short selling against the box was prevalent prior to TRA97, using relative short interest as our dependent variable.[14] We employ several independent variables in our regression analysis. The larger is the stock price appreciation in the tax year to date, the larger the incentive to postpone taxes by shorting against the box; we therefore include the stock return from January 1 to the 15th of the current month (corresponding with the settlement date for short interest reported by both NYSE and NASDAQ). Short selling is likely to be more attractive when traders are able to cover their short position easily. Thus, stocks with high trading volume are more liquid and more likely to be sold short, ceteris paribus. Therefore, we include in our regressions the natural log of average daily volume from the 15th of the previous month to the 15th of the current month.

A strong correlation is found between firm size and institutional ownership (see Sias and Starks 1997). As such, stocks with larger market capitalization are likely to have a relatively high availability of shares to be sold short. We control for the availability of shares by including the

14. The regression results we obtain are qualitatively similar when we use alternate measures of short selling such as total short interest, natural log of short interest, natural log of relative short interest, and short interest divided by average daily trading volume (not reported). Statistical tests (see Hsiao 1986, pp. 12–15) indicate that our cross-sectional time series data can be treated as pooled data. Consequently, all our regressions are pooled ordinary least squares regressions, with appropriate robust standard errors as discussed in the table captions.

market value of the firm in our regressions. Market value is calculated as the natural logarithm of the product of the stock price at the market close on the 15th of the month and the shares outstanding as of the 15th of the month.

Shares may be sold short to hedge option positions in the stock. To control for options arbitrage and hedging-motivated short interest, we construct a dummy variable to indicate whether the stock has traded options. We find information on which stocks have traded options from Options Clearing Corporation. The option dummy variable takes a value of 1 if a stock has traded options (264 of 535 NYSE securities, 504 of 1,129 NASDAQ securities), and 0 otherwise. In addition to options arbitrage and hedging effects on short interest, the option market may provide a substitute for short selling through synthetic construction of short positions (see note 7 though). We also control for the possibility of arbitrage-motivated short selling in conjunction with the presence of convertible debt or convertible preferred stock. Research Insight (Compustat) provides data on the firms in our sample that have convertible securities. We construct a dummy variable that takes a value of 1 if a company has convertible securities in a given month and 0 otherwise. Because the data are reported on an annual basis, we may lose some power in our tests. Nonetheless, we find strong statistical significance in the direction that we expect.

To capture the effect of shorting due to merger-related risk arbitrage, we include a dummy variable in our regression to control for merger arbitrage activity. The dummy variable is set to 1 for any firm that announces an acquisition and remains 1 until the acquisition is completed and the target delists.[15] Our merger data correspond to acquisition announcements that involve stock as part of the compensation. Both the acquirer and the target must be publicly traded. Within our data, we find 187 NYSE acquisitions (92 using a fixed stock ratio and 95 a floating stock ratio), comprising 86 different firms. The average time to completion is 6.2 months. We find 185 NASDAQ acquisitions (124 using a fixed stock ratio and 61 a floating stock ratio), comprising 121 different firms. The average time to completion is 6.0 months. Six of the NASDAQ acquisitions and 13 of the NYSE acquisitions subsequently fell through; we treated these as 6-month deals so that we do not lose the data points. Our results are not sensitive to excluding the 19 mergers that fell through.

Movements of the overall market may affect traders' ability or desire to sell stocks short. To control for this, we include in our regression analysis the return on the CRSP value-weighted index calculated between the 15th of the previous month and the 15th of the current month. Similarly, we include a measure of aggregate "market" short interest to

15. These data were kindly provided by Todd Pulvino. See Mitchell et al. (2003) for details on data collection.

capture systematic changes and trends in short interest. We compute this variable by summing the short interest for all the NYSE stocks in our sample for the NYSE regressions and by summing the short interest for all the NASDAQ stocks in our sample each month.

We are interested in the effects of TRA97 on short selling and on the predictive power of short-interest announcements and are particularly interested in short-selling activity in December and January before and after the implementation of TRA97. We use dummy variables to capture unique aspects of these time periods. BDEC is a dummy variable that is set to 1 for data that are in the month of December and before June 8, 1997; otherwise, it is 0. Thus, it captures any temporal effects on short selling due to Decembers pre TRA97. ADEC is a dummy variable that is set to 1 for data that are in the month of December and after June 8, 1997; otherwise, it is 0. BJAN is a dummy variable that is set to 1 for data that are in the month of January and prior to June 8, 1997; otherwise, it is 0. AJAN is a dummy variable that is set to 1 for data that are in the month of January and after June 8, 1997; otherwise, it is 0. By examining the differences among the coefficients for BDEC, BJAN, ADEC, and AJAN, we can assess the extent to which market participants' behavior is consistent with the proposed "minimum-cost strategy" and we can deduce the impact of TRA97 on short selling against the box.

The independent variables in the following regression control for short interest determinants and allow us to identify short selling that we believe to be short selling against the box; we run the following regression on NYSE stocks and NASDAQ stocks:

$$\begin{aligned}\text{Relative short interest} = {} & b_0 + b_1 * (\text{aggregate short interest}) \\ & + b_2 * (\text{stock return}) + b_3 * \text{BDEC} + b_4 * \text{BJAN} \\ & + b_5 * \text{ADEC} + b_6 * \text{AJAN} + b_7 * (\text{volume}) \\ & + b_8 * (\text{firm size}) + b_9 * (\text{option dummy}) \\ & + b_{10} * (\text{market index return}) \\ & + b_{11} * (\text{convertible dummy}) \\ & + b_{12} * (\text{merger dummy}) + e. \qquad (1)\end{aligned}$$

Table 6 reports the results of regression 1. Aggregate short interest and dummy variables for options, convertibles, and merger arbitrage are all positively related to short-selling activity and consistently statistically significant for both exchanges.

Firms with traded options, convertible securities, or those that announce acquisition plans have more relative short interest than their counterparts, ceteris paribus. These differences are likely to be largely due to hedging and arbitrage activities. More specifically, NYSE firms with traded options have on average 1.4 percentage points more relative

short interest than those without traded options (0.7 percentage points for NASDAQ firms); NYSE firms with convertible securities have 1.3 percentage points more relative short interest than those without convertibles (1.3 percentage points for NASDAQ firms); and NYSE firms with ongoing merger activity have 1.6 percentage points more relative short interest than those without (0.8 percentage points for NASDAQ firms). As a basis for comparison, our NYSE firms have an average relative short interest of 1.34% and standard deviation of 3.03%, and our NASDAQ firms have average relative short interest of 1.70% with standard deviation of 3.24%. Thus, options, convertibles, and mergers all have a substantial impact on relative short interest, and collectively, these findings highlight the importance of arbitrage and hedging activities on relative short-interest levels.

Volume and size proxy for the liquidity and availability of the stock; positive relationships between these variables and short-interest would be consistent with previous research. We find this for volume, but we find a negative coefficient on size. Although size and volume are highly correlated (see table 5), this is not a multicollinearity problem: performing the regression with (1) size only, (2) volume only, and (3) both volume and size for both NYSE and NASDAQ securities yields similar results. Rather, the negative coefficient on size in table 6 is a direct and mechanical result of having relative short interest (which is shares held short divided by shares outstanding, a measure closely related to size) on the left-hand side. Reformulating the regression with log of short interest on the left-hand side, for example, gives the expected signs on all coefficients (and higher R^2). The results are qualitatively very similar to the results presented here, but we think the coefficients are easier to interpret as we have presented them, with relative short interest on the left-hand side.

We examine the differences between BJAN and BDEC and between ADEC and BDEC to assess whether there was a change in short-selling activity after the implementation of TRA97. Specifically, if short selling against the box had been prevalent before TRA97 and uncommon after, then prior to TRA97, short interest would have been relatively large in December (as some traders deferred capital gains taxes) and relatively small in January (as they closed their costly positions), but there would be essentially no difference between December and January short selling after TRA97. Following Dyl (1978), we hypothesize that the following relation among regression coefficients holds: $\text{BJAN} - \text{BDEC} < 0$. After TRA97, we anticipate no significant difference between January and December short interest because the end-of-year tax motivations for opening or closing short positions is gone. So, we hypothesize that the following relation among regression coefficients holds: $\text{AJAN} - \text{ADEC} = 0$.

Tests for differences among coefficients are reported in panel B of Table 6. The results are consistent with our expectations. The differences

TABLE 6 Determinants of Short Interest

A. Regression Results (Dependent Variable = Relative Short Interest)

	NYSE	NASDAQ
Intercept	.012232	−.058445
	(2.88***)	(−16.20***)
Aggregate short interest (÷1,000)	.006671	.003001
	(2.60***)	(1.87*)
Stock return (YTD from Jan. 1)	.004347	−.001250
	(1.82*)	(−1.54)
Dummy: December before TRA97 (BDEC)	−.001216	.000172
	(−2.16**)	(.36)
Dummy: January before TRA97 (BJAN)	−.002765	−.003666
	(−4.59***)	(−9.71***)
Dummy: December after TRA97 (ADEC)	.001175	.000730
	(1.55)	(1.40)
Dummy: January after TRA97 (AJAN)	.000918	−.000209
	(1.46)	(−.43)
Log natural (average daily trading volume)	.006577	.006807
	(9.55***)	(23.43***)
Log natural (market value of equity)	−.006385	−.000394
	(−8.17***)	(−1.33)
Dummy: traded options	.014255	.007118
	(9.87***)	(7.51***)
Market return	−.004414	−.004852
	(−2.09**)	(−3.48***)
Dummy: convertible securities	.013005	.013408
	(5.57***)	(9.08***)
Dummy: merger arbitrage	.016146	.008359
	(4.56***)	(3.29***)
Adjusted R^2	.1647	.1880
F-statistic	510.88***	1,264.70***

B. Difference Tests

Coefficients	Expected Sign	Intuition	NYSE Difference (*t*-Statistic)	NASDAQ Difference (*t*-Statistic)
BJAN–BDEC	Negative	Before TRA97, short selling against the box drives up short interest before turn of the year; traders close positions in January	−.00155 (−2.16**)	−.00384 (−6.71***)
AJAN–ADEC	Zero	After TRA97, no tax motivations to short sell, so no calendar effects expected	−.00026 (−.35)	−.00094 (−1.37)

NOTE.—Data are monthly observations for 535 NYSE stocks and 1,129 NASDAQ stocks during the period January 1995 through November 1999. The dependent variable is relative short interest (short interest divided by number of shares outstanding). The independent variables are an intercept, aggregate short interest for all stocks in our sample divided by 1,000 (NYSE only for the NYSE regressions, NASDAQ only for the NASDAQ regressions), the individual stocks' tax-year-to-date return up to that month's short interest announcement date (i.e., the rate of return from January 1 to the 15th of the month of the short interest announcement), dummy variables to indicate December before TRA97, December after TRA97, January before TRA97, and January after TRA97, the natural logarithm of 1 plus the average daily volume over the previous month, the log of 1 plus the firm market value, a dummy variable indicating if there are traded options on the stock, the monthly return on the CRSP value-weighted index, a dummy variable to indicate if the company has convertible securities outstanding, and a merger arbitrage dummy variable that takes the value 1 if the firm has announced an acquisition of another publicly traded firm (we change this to 0 when the target delists). All *t*-statistics are corrected for heteroscedasticity and autocorrelation using 12 months of lags (see Newey and West 1987) and are reported in parentheses. *, **, ***denote significance at the 10%, 5%, and 1% levels, respectively.

between Decembers and Januarys in the pre-TRA97 period are about 16 basis points of relative short interest for NYSE stocks in our sample and about 38 basis points of relative short interest for NASDAQ stocks in our sample. These differences are statistically different from zero (*t*-statistics of −2.17 and −6.71 for NYSE and NASDAQ, respectively). Economically, this represents about 11.6% (39.2%) of mean (median) short interest on NYSE and 22.5% (81.2%) of mean (median) short interest on NASDAQ (see tables 3 and 4 for means and medians). These means and medians are not themselves very large, but our results indicate that short selling against the box had a significant effect on what short selling there was prior to TRA97.[16]

VI. Noise Reduction: Increasing the Predictive Power of Short Interest

This section contains our central findings. Based on our theoretical results from Section III, our main hypothesis is that TRA97 increased the bearish information content of short interest in stocks. In addition to some of the variables defined previously, we employ the following variables in our regression that examines the information content of short interest announcements. We use two different model specifications to differentiate between levels of relative short interest. First, we use a continuous variable to measure short interest: we use the variables BSISHR and ASISHR to denote relative short interest before and after TRA97, respectively.

To allow a less restrictive specification, we also regress returns on quintile dummy variables.[17] These are dummy variables that are set to 1 if an observation for a stock in a given month is in that quintile of relative short interest (a preceding *B* denotes being in that quintile in a month that is before TRA97; a preceding *A* denotes being in that quintile in a month that is after TRA97). Otherwise, the variable is set to 0. The variables BQ5 and AQ5 denote the highest quintile of relative short interest before and after TRA97, respectively. These variables allow us to examine whether "extreme" short interest conveys more information than moderate levels of short interest, and they do not impose the constraint that the effect must be the same in all quintiles.

We report the results of regressions 2 and 3 in table 7:

$$\begin{aligned}\text{Future 1-month stock return} = {} & b_0 + b_1 * (\text{size}) \\ & + b_2 * (\text{future 1-month market return}) \\ & + b_3 * \text{BSISHR} + b_4 * \text{ASISHR} + e, \qquad (2)\end{aligned}$$

16. In our sample we have only two Januarys after TRA97, January 1998 and January 1999. Other things being equal, this tends to reduce the power of the second test in panel B of table 4 (i.e., the one with null that there were no tax-induced calendar effects on short interest after TRA97). However, our cross section is so large that issues of power should not be a concern.

17. We also test specifications using quartiles and deciles, which yield similar results.

TABLE 7 Short Interest and Future Return

A. Regression Results (Dependent Variable = 1-Month Future Return)

	NYSE		NASDAQ	
Intercept	−.0526	−.0460	−.1149	−.1078
	(−11.32***)	(−11.22***)	(−19.72***)	(−19.25***)
Log normal (market value of equity)	.0041	.0032	.0103	.0090
	(11.17***)	(10.59***)	(20.88***)	(19.69***)
Market index return	.7663	.7748	1.2376	1.2406
	(49.30***)	(49.73***)	(62.23***)	(62.84***)
Relative short interest pre TRA97 (BSISHR)		.0616		−.2319
		(2.11**)		(−7.61***)
Relative short interest post TRA97 (ASISHR)		−.1390		−.1782
		(−3.34***)		(−5.40***)
Relative short interest quintile 2 pre TRA97 (BQ2)	.0022		.0025	
	(1.31)		(1.02)	
Relative short interest quintile 2 post TRA97 (AQ2)	−.0124		−.0105	
	(−6.22***)		(−3.74***)	
Relative short interest quintile 3 pre TRA97 (BQ3)	−.0003		−.0092	
	(−.15)		(−3.87***)	
Relative short interest quintile 3 post TRA97 (AQ3)	−.0137		−.0171	
	(−5.85***)		(−6.33***)	
Relative short interest quintile 4 pre TRA97 (BQ4)	−.0019		−.0181	
	(−.99)		(−7.17***)	
Relative short interest quintile 4 post TRA97 (AQ4)	−.0136		−.0203	
	(−5.54***)		(−7.06***)	
Relative short interest quintile 5 pre TRA97 (BQ5)	−.0025		−.0274	
	(−1.14)		(−10.72***)	
Relative short interest quintile 5 post TRA97 (AQ5)	−.0168		−.0179	
	(−6.38***)		(−5.78***)	
Adjusted R^2	.0974	.0944	.0733	.0721
F-statistic	335.7***	809.2***	518.1***	1273.5***

TABLE 7 (*Continued*)

B. Difference Tests

Coefficients	Expected Sign	NYSE Difference (*t*-Statistic)	NASDAQ Difference (*t*-Statistic)
ASISHR – BSISHR	Negative	−.2006 (−4.17***)	.0537 (1.29)
AQ5 – BQ5	Negative	−.0143 (−4.98***)	.0095 (2.85***)
AQ4 – BQ4	Negative	−.0117 (−4.97***)	−.0023 (−.75)
AQ3 – BQ3	Negative	−.0134 (−5.85***)	−.0079 (−2.68***)
AQ2 – BQ2	Negative	−.0146 (−6.90***)	−.0130 (−4.13***)
Wald test statistic†		131.41***	32.90***

NOTE.—Data are monthly observations for 535 NYSE and 1,129 NASDAQ stocks during the period January 1995 through November 1999. The dependent variable is the individual stocks' 1-month future return. The independent variables are an intercept, firm size as measured by the natural logarithm of the market value of equity, the market return contemporaneous with the stock's return as measured by the 1-month return on the CRSP value-weighted index. Returns are alternately calculated based on 1 month subsequent to the announcement of short interest. Other independent variables are relative short interest in the stock before or after TRA97, dummy variables indicating if the quintile of relative short interest for the stock (Quintile 5 is highest short interest) before and after TRA97. All *t*-statistics are corrected for heteroscedasticity (see White 1980) and reported in parentheses. The results are robust to autocorrelation (robustness tests using corrections for autocorrelation with 3-, 6-, and 12-month lags give virtually identical results to those reported here, and are omitted (see Newey and West 1987). The results of difference tests among the coefficients and a Wald test for joint significance are reported in panel B. *, **, ***denote significance at the 10%, 5%, and 1% levels, respectively. †denotes a joint test that all quintile differences are zero.

$$\begin{aligned}\text{Future 1-month stock return} = {} & b_0 + b_1 * (\text{size}) \\ & + b_2 * (\text{future 1-month market return}) \\ & + b_3 * \text{BQ2} + b_4 * \text{BQ3} + b_5 * \text{BQ4} \\ & + b_6 * \text{BQ5} + b_7 * \text{AQ2} + b_8 * \text{AQ3} \\ & + b_9 * \text{AQ4} + b_{10} * \text{AQ5} + e, \end{aligned} \tag{3}$$

where in addition to the previously defined variables, we control for market movements with the 1-month return on the market (CRSP value-weighted return contemporaneous with the future stock return). The dependent variable is the 1-month future return on the stock. We use two different time frames for calculation of returns. First, we use the 1-month period from the short interest announcement to the following month's announcement of short interest. The announcement reflects the first time that short interest information is readily available to all market participants. It is possible, however, that some market participants might know the short interest information before the public announcement and be able to trade based on that information; so for

robustness, we also calculate returns based on when short interest is calculated by the exchanges (the settlement date), which is several days earlier (4 days for NYSE, 8 days for NASDAQ). The results are qualitatively very similar, so we report only those results based on the news release dates.

We test for statistical differences between coefficients to determine if the informational content of short interest has increased. Specifically, we test for differences between the regression coefficients on the following variables: ASISHR and BSISHR, AQ2 and BQ2, AQ3 and BQ3, AQ4 and BQ4, and AQ5 and BQ5.[18] Our extension to Diamond and Verrecchia's model implies that the informativeness of short interest should increase with the exit of tax-deferring short sellers against the box. Therefore, we expect the coefficients of the quintile variables and continuous short interest variable to become more negative after TRA97, making the expected sign on all of our difference tests negative.

The results of these regressions are in table 7, panels A and B. Prior to TRA97, high levels of short interest (represented by BQ5) have a significantly negative relationship with future returns for NASDAQ stocks and an insignificantly negative relationship with future returns for NYSE stocks. Lower levels (represented by BQ2) have an insignificantly positive relationship with returns. This relationship is consistent with Asquith and Meulbroek (1995), short interest is bearish in the most heavily shorted stocks.

The most important information from our regressions, though, is not the values of the coefficients themselves but the differences in coefficient values. Panel B of table 7 reports the results of difference tests among our temporal dummy variables and continuous measures of levels of relative short interest. For the NYSE stocks, short-interest announcements convey significantly more negative information after TRA97. Difference tests for our quintile dummy variables and our relative short-interest variable are negative, as expected, and statistically significant for both news release date and settlement date returns. The differences for the dummy variables are of a magnitude of about negative 1.3% on average. Our Wald test statistic strongly rejects that the differences of the coefficients are jointly zero. After TRA97, short-interest announcements thus convey substantially more negative information for NYSE stocks.

The results differ for NASDAQ stocks. Short interest conveys considerably more negative information for heavily shorted NASDAQ stocks than heavily shorted NYSE stocks before TRA97. Short interest in

18. Although not reported, we examined the robustness of our results to using other measures of short interest: short interest divided by average daily trading volume, log of short interest, change in relative short interest, and percent change in short interest all provide qualitatively similar results.

the top quintile of NASDAQ short interest is higher than the short interest in the corresponding NYSE quintile. If a high level of negative information leads to only a small marginal impact of an increase in costs, then we should not expect TRA97 to have as much impact on our NASDAQ stocks. Empirically, we find that, for our relative short-interest variable and the two highest quintiles of relative short interest, the differences in coefficients in our NASDAQ regressions do not support our hypothesis that the information content of short interest increases after TRA97. The differences are positive (and significant for the most heavily shorted stocks) or only weakly negative. However, evidence from more moderate levels of relative short interest (Quintiles 2 and 3) do support Diamond and Verrecchia's model—the differences are negative and statistically significant. We suggest a possible explanation for this NASDAQ result.

NASDAQ underwent substantial market reforms throughout 1997. Due to these reforms, dealers faced increased competition from limit orders and electronic communication networks (ECNs). These reforms dramatically decreased trading costs on the NASDAQ (see Weston 2000). Decreases in trading costs make trading more attractive for uninformed traders, and an influx of uninformed traders in response to the market reforms may have more than offset the noise-reducing effects that TRA97 had on the information content of short interest. In addition to the significant structural shifts around the time of TRA97 (see Weston 2000), during our sample period, there was an increase in electronic trading and day trading, much of it focused on NASDAQ stocks (see Garvey 2001). Many of these day traders were found to be short selling without regard to the usual restrictions and costs associated with short selling (see Arnold and Butler 2003). In addition to the structural changes and the changes in the composition of traders, during the "tech bubble," the NASDAQ index posted unusually high annual returns of approximately 40% in 1998 and over 80% in 1999. Collectively, these changes in the NASDAQ market during our sample period mean that our NASDAQ results must be interpreted with caution.

VII. Conclusions and Directions for Further Research

As indicated in figure 3, short selling increased over the period 1995–98. Short selling against the box dropped off dramatically following the implementation of the Taxpayer Relief Act of 1997, however. With the absence of investors selling short against the box, the informativeness and predictive ability of short interest materially increased. This provides strong support for the predictions of Diamond and Verrecchia (1987).

These conclusions are important: the exit from the market of tax-driven, uninformed short sales against the box means that the remaining

short positions convey more information than before the Tax Relief Act of 1997. That is, short sales are now more likely to be informative due to the "noise reduction" of eliminating uninformed, tax-motivated short selling against the box.

Whereas we focus on only one facet of the Taxpayer Relief Act of 1997 (increased cost of short selling), the act also changed capital gains tax rates, eliminated tax deferring benefits to synthetic option positions (including some collars), and changed the holding periods for long- versus short-term gains. Thus, TRA97 might provide a useful laboratory for researchers studying other financial phenomena.

Appendix

This appendix derives and proves our two main theoretical results. Using our extended version of the Diamond and Verrecchia (1987) model, we calculate the proportion of informed short selling by dividing the probability of informed short selling by the probability of short selling.

From the game tree, the probability of uninformed short selling is

$$\frac{1}{2}g(1-a)[h_q(c_1+c_2)+(1-h)c_1], \tag{A1}$$

and the probability of informed short selling is

$$\frac{1}{2}ga(c_1+c_2)[(1-h)+hq]. \tag{A2}$$

We can determine the proportion of informed short selling by dividing equation (A1) by the sum of equations (A1) and (A2). Thus, the proportion of informed short selling is

$$\frac{a(c_1+c_2)}{(c_1+ac_2)}\left\{\frac{(1-h)+hq}{(1-h)+hq\left[\dfrac{c_1+c_2}{c_1+ac_2}\right]}\right\}. \tag{A3}$$

We now derive our main theoretical results.

PROPOSITION 1. As the probability of short selling against the box, q, decreases, the proportion of informed short selling increases.

Proof. First, note that, because $0 < a < 1$, $(c_1+c_2)/(c_1+ac_2) > 1$. Now, it is easy to see that the partial derivative of equation (A3) with respect to q is negative:

$$\frac{a(c_1+c_2)}{(c_1+ac_2)}\left\{\frac{h(1-h)\left[1-\left(\dfrac{c_1+c_2}{c_1+ac_2}\right)\right]}{\left[(1-h)+hq\left(\dfrac{c_1+c_2}{c_1+ac_2}\right)\right]^2}\right\}<0 \tag{A4}$$

Q.E.D.

Corollary to Proposition 1. If q is zero, the proportion of informed short selling is greater than if q is strictly positive.

Proof. We first establish the following result.

Lemma. Equation (A3) is less than $a(c_1 + c_2)/(c_1 + ac_2)$.

Proof. Assume $0 \leq a < 1$; that is, not all traders are informed, $c_1 > 0$, and $c_2 > 0$. Then,

$$\left[\frac{c_1 + c_2}{c_1 + ac_2}\right] > 1. \tag{A5}$$

Therefore,

$$(1 - h) + hq\left[\frac{c_1 + c_2}{c_1 + ac_2}\right] > (1 - h) + hq. \tag{A6}$$

Therefore,

$$\left[\frac{(1 - h) + hq}{(1 - h) + hq\left(\frac{c_1 + c_2}{c_1 + ac_2}\right)}\right] < 1. \tag{A7}$$

Therefore,

$$\frac{a(c_1 + c_2)}{(c_1 + ac_2)}\left[\frac{(1 - h) + hq}{(1 - h) + hq\left(\frac{c_1 + c_2}{c_1 + ac_2}\right)}\right] < \frac{a(c_1 + c_2)}{(c_1 + ac_2)}. \tag{A8}$$

Q.E.D.

We now prove the corollary by setting $q = 0$ in equation (A3). Equation (A3) collapses to

$$\frac{a(c_1 + c_2)}{(c_1 + ac_2)}. \tag{A9}$$

Hence, by the lemma, the corollary is true.
Q.E.D.

References

Aitken, Michael J., Alex Frino, Michael S. McCorry, and Peter L. Swan. 1998. Short sales are almost instantaneously bad news: Evidence from the Australian stock exchange. *Journal of Finance* 53:2205–23.

Alexander, Gordon J., and Mark A. Peterson. 1999. Short selling on the New York Stock Exchange and the effects of the uptick rule. *Journal of Financial Intermediation* 8, 90–116.

Angel, James J. 1997. Short selling on the NYSE. Working paper, Georgetown University.

Arnold, Tom, and Alexander W. Butler. 2003. The informativeness of short sales during a period of uptick rule violations. Working paper, University of Richmond and University of Texas at Dallas.

Asquith, Paul, and Lisa K. Meulbroek. 1995. An empirical investigation of short interest. Working paper, Massachusetts Institute of Technology and Harvard University.

Bessembinder, Hendrik. 2000. Tick size, spreads, and liquidity: An analysis of NASDAQ securities trading near ten dollars. *Journal of Financial Intermediation* 9:213–39.

Brent, Averil, D. Morse, and E. Kay Stice. 1990. Short interest: Explanations and tests. *Journal of Financial and Quantitative Analysis* 26:273–89.

Cornelli, F., and D. Li. 2002. Risk arbitrage in takeovers. *Review of Financial Studies* 15: 837–68.

Danielson, Bartley R., and Sorin M. Sorescu. 2001. Why do option introductions depress stock prices? A study of diminishing short-sale constraints. *Journal of Financial and Quantitative Analysis* 36:451–84.

D'Avolio, Gene. 2002. The market for borrowing stock. *Journal of Financial Economics* 66:271–306.

Dechow, Patricia M., Amy P. Hutton, L. Meulbroek, and Richard G. Sloan. 2001. Short-sellers, fundamental analysis and stock returns. *Journal of Financial Economics* 61:77–106.

Desai, Hemang, K. Ramesh, S. R. Thiagarajan, and Bala Balachandran. 2002. An investigation of the informational role of short interest in the NASDAQ market. *Journal of Finance* 57:2263–87.

Diamond, Douglas W., and Robert E. Verrecchia. 1987. Constraints on short-selling and asset price adjustment to private information. *Journal of Financial Economics* 18:277–311.

Dyl, E. 1978. Short selling and the capital gains tax. *Financial Analysts Journal* 34:61–64.

Figlewski, Stephen, and Gwendolyn P. Webb. 1993. Options, short sales, and market completeness. *Journal of Finance* 48:761–77.

Garvey, Ryan. 2001. Do day traders profit from NASDAQ market makers? An analysis of day trading profits. Working paper, University College Dublin.

Goldstein, M. A., and K. A. Kavajecz. 2000. Eights, sixteenths and market depth: changes in tick size and liquidity provision on the NYSE. *Journal of Financial Economics* 56:125–49.

Graham, J. Edward, J. Christopher Hughen, and Cynthia G. McDonald. 1999. The determinants of short sales activity. Working paper, University of Missouri—Columbia.

Hsiao, Cheng. 1986. *Analysis of panel data*. New York: Cambridge University Press.

Jones, Charles M., and Owen A. Lamont. 2002. Short sale constraints and stock returns. *Journal of Financial Economics* 66:207–39.

Macaulay, Fred R., and David Durand. 1951. Short selling on the New York Stock Exchange. New York: Twentieth Century Fund.

Miller, Edward M. 1977. Risk, uncertainty, and divergence of opinion. *Journal of Finance* 32:1151–68.

Mitchell, M., and T. Pulvino. 2001. Characteristics of risk in risk arbitrage. *Journal of Finance* 56:2135–76.

Mitchell, Mark, Todd Pulvino, and Erik Stafford. 2004. Price pressure around mergers. *Journal of Finance* 59:31–63.

Newey, W., and K. West. 1987. A simple positive semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55:703–08.

Reed, Adam. 2002. Costly short-selling and stock price adjustment to earnings announcements. Working paper, University of North Carolina.

Senchack, A. J., and Laura T. Starks. 1993. Short-sale restrictions and market reaction to short-interest announcements. *Journal of Financial and Quantitative Analysis* 28:177–94.

Sias, Richard W., and Laura T. Starks. 1997. Institutions and individuals at the turn-of-the-year. *Journal of Finance* 52:1543–62.

Tucker, Lawrence E., and Mark T. Watson. 1999. Short sales against the box after the TRA '97. *Tax Adviser* 30:246–48.

Weston, James P. 2000. Competition on the NASDAQ and the impact of recent market reforms. *Journal of Finance* 55:2565–98.

White, Halbert. 1980. A heteroskedasticity-consistent covariance estimator and a direct test for heteroskedasticity. *Econometrica* 48:817–38.