THE JOURNAL OF FINANCE • VOL. LIV, NO. 6 • DECEMBER 1999

The Delisting Bias in CRSP's Nasdaq Data and Its Implications for the Size Effect

TYLER SHUMWAY and VINCENT A. WARTHER*

ABSTRACT

We investigate the bias in CRSP's Nasdaq data due to missing returns for delisted stocks. We find that the missing returns are large and negative on average, and that delisted stocks experience a substantial decrease in liquidity. We estimate that using a corrected return of −55 percent for missing performance-related delisting returns corrects the bias. We revisit previous work which finds a size effect among Nasdaq stocks. After correcting for the delisting bias, there is no evidence that there ever was a size effect on Nasdaq. Our results are inconsistent with most risk-based explanations of the size effect.

EACH YEAR, MANY OF THE STOCKS TRADED on Nasdaq are delisted from the system and cease to be traded there. Delistings occur for a number of reasons including merger and acquisition, bankruptcy, liquidation, or migration to another exchange. They usually coincide with significant firm-specific events, so the returns associated with delistings are often sizable. Further, a stock's liquidity can change dramatically upon delisting, affecting the price at which shareholders can sell their shares. The Center for Research in Security Prices (CRSP) attempts to systematically collect delisting returns; however, the task is difficult, and many returns are not collected. Some categories are missed more often than others. Specifically, delisting returns associated with poor firm performance (e.g., bankruptcy or failure to meet capital requirements) are missed much more often than returns associated with neutral or good firm performance (e.g., merger, acquisition, or migration to another exchange). Since most of the missing delisting returns are associated with negative events, a significant bias exists in the data.

We study this delisting bias in two ways. First, we document the magnitude of the bias. We investigate when and how often delisting returns are not collected by CRSP, and we estimate the average size of the missing returns. In this way, we extend to Nasdaq the work of Shumway (1997) which examines the delisting bias in CRSP's NYSE and AMEX data. The bias in the Nasdaq data is much larger—on average, 1.2 percent of NYSE and AMEX stocks are delisted for poor performance each year but 5.6 percent of Nasdaq stocks are delisted each year for similar reasons. Delistings are most fre-

* University of Michigan Business School. We thank Dmitry Davydov for superb research assistance. We also thank Gene Fama, Mark Mitchell, René Stulz, an anonymous referee, and participants at a seminar at the University of Michigan for helpful comments.

quent for the smallest Nasdaq stocks—on average, 2.95 percent of stocks in the smallest five percent of Nasdaq stocks are delisted each month. Using over-the-counter prices taken from the *Pink Sheets*, we locate return data for 63 percent of these missing returns. The average return for the other 37 percent is presumably even lower. We also document a substantial decrease in liquidity upon delisting which further decreases the effective value of the stock. Accounting for all these factors, we estimate the average effective return for stocks delisted from Nasdaq for performance reasons to be −55 percent. Researchers can use this estimated return to correct the CRSP database for the delisting bias.

Second, we explore the implications of this delisting bias for the size effect first documented by Banz (1981) and Reinganum (1981). We revisit the work of Lamoureux and Sanger (1989) which finds evidence of a size effect in the Nasdaq data, and we find that the results are sensitive to the delisting bias since delistings occur more frequently among smaller stocks. After correcting for the bias, virtually no evidence of a size effect remains for the full sample period from 1972 to 1995. We repeat the analysis for the two subperiods covering the years from 1973 to 1985 and from 1986 to 1995, and we find no evidence of a size effect in either subperiod. Hence, there is no evidence that there ever was a size effect among Nasdaq stocks.

The results are relevant to the debate over the interpretation of the size effect. Many of the explanations proposed for the size effect assert that small firms earn high average returns because they are fundamentally riskier than large firms. For example, Berk (1995) argues that size proxies for unmeasured systematic risk, and both Chan and Chen (1991) and Fama and French (1995) contend that the higher returns are compensation for the higher risk of financial distress among small firms. Dichev (1998), however, finds evidence inconsistent with the risk explanation of the size effect.

Small Nasdaq stocks are ideal for examining the size effect because, as a group, they are the smallest and most distressed stocks for which we have reliable data. Presumably, the size effect should be most pronounced for this sample of stocks. Using delisting frequency as a proxy for distress, we show that size and distress are correlated among Nasdaq stocks. Consequently, our finding of no size effect among Nasdaq stocks is evidence against the risk-based explanations of the size effect. This is consistent with Dichev (1998) as well as the assertions of Lo and MacKinlay (1990) and Black (1993) that the size effect is an example of a data-snooping bias rather than a fundamental economic phenomenon.

The delisting bias also has implications for studies that compare the average returns of NYSE and Nasdaq stocks. Reinganum (1990) first reported that small NYSE firms have average annual returns about six percent higher than Nasdaq firms of the same size. Loughran (1993) attributes about 60 percent of the difference to the poor performance of initial public offerings; Fama et al. (1993) attribute the difference to a higher probability of distress for NYSE firms. Correcting for the delisting bias documented here increases the

return differential further since the delisting bias is larger in the Nasdaq data. This increases the return-differential puzzle at the center of this literature.

The paper proceeds as follows. Section I documents the delisting bias in Nasdaq data, Section II estimates an average return that researchers can use to correct for the bias, Section III analyzes the implications of the bias for the size effect, and Section IV concludes.

I. The Delisting Bias in CRSP's Nasdaq Data

A. The Frequency of Missing Delisting Returns

To assess the frequency with which delisting returns are omitted from CRSP, we examine daily returns for all the stocks in CRSP's Nasdaq database from 1972 to 1995. Whenever a stock is coded as having been delisted, we record whether a delisting return is given.[1] We classify delisting codes 500 and 505 to 588 as performance-related (i.e., a negative change for the firm). We do not classify delisting codes 501 and 502 (migration to the NYSE or AMEX) as performance-related because these events are usually positive changes for a firm (see Baker and Edelman (1991)). Table I shows the number of delistings for each category established by CRSP, and performance-related reasons are the most common given for delisting, followed by merger and migration to the NYSE or AMEX.[2]

No reason is given for 26 percent of the performance-related delistings. This designation exists largely because prior to 1987 CRSP assigned all distressed delistings a code of 5, so the events in this category occur for the same reasons as the other events in the table. Since most of the others are performance-related, we classify all of the "reason unavailable" delists as performance-related.

The pattern of missing returns is striking. The three nonperformance-related categories have virtually no missing returns. No more than one percent of returns are missing for merger, exchange, or movement to the NYSE or AMEX. However, virtually all (99.8 percent) returns are missing for performance-related delistings. Hence, the missing returns will almost certainly introduce a bias. The much smaller category of liquidations is a special case, for although 71.0 percent are missing returns, CRSP gives them a code indicating the liquidation was announced before the delisting. Conse-

[1] In the 1996 CRSP monthly files, CRSP has added partial-month returns to the delisting return field, which are calculated using the last recorded month-end price and the last day price of each delisted security. While CRSP calls these returns "delisting returns," they should not be confused with the delisting returns we examine here. CRSP's returns use only predelisting prices. Our delisting returns use postdelisting prices and so include the effect of the delisting.

[2] Although there are relatively few bankruptcy-related delistings recorded by CRSP, many of the firms delisted for other reasons are already bankrupt when they delist. It is common, for example, for a firm to file for bankruptcy and then be delisted for failing to meet Nasdaq's financial guidelines.

Table I

Delisting Returns in CRSP's Nasdaq Data

Delistings are categorized according to the reason given by CRSP for the delisting. Shown are the number of delistings that fall in each category and the percentage of returns in each category for which the delisting return (DRET) is missing. Data are from CRSP's 1995 daily Nasdaq file. The sample period is 1972 through 1995.

Reason for Delisting	CRSP Code	Number of Delistings	Percentage of All Performance Delistings	Percentage Missing DRET
Merger	200–240	2463		0.9
Exchange	300–390	292		1.0
Liquidation	400	138		71.0
Moved to NYSE or AMEX	501, 502	1476		0.7
Performance-related delistings				
Reason unavailable	500	1087	26.0	
Insufficient capital, surplus, equity	560	987	23.6	
Delinquent in filing or fees	580	652	15.6	
Insufficient market makers	550	470	11.3	
Insufficient float or assets	561	241	5.8	
Price too low	552	174	4.2	
Company request (no reason)	570	155	3.7	
Insufficient shareholders	551	98	2.3	
Failure to register under 12G of SEA	581	71	1.7	
Moved to *Pink Sheets*	520	61	1.5	
Moved to Boston Exchange	510	34	0.8	
Bankruptcy or insolvency	574	27	0.6	
Company request (liquidation)	572	23	0.6	
Failure to meet equity requirements	582	20	0.5	
Company request (private)	573	18	0.4	
Moved to Pacific Exchange	516	13	0.3	
Delisted by the SEC	700	13	0.3	
Does not meet financial guidelines	584	12	0.3	
Issue withdrawn	575	11	0.3	
Protection of investors/public	585	8	0.2	
Denied temporary exception	583	6	0.1	
Moved to Philadelphia Exchange	517	3	0.1	
Corporate governance violations	587	2	0.0	
Moved to Toronto Exchange	519	1	0.0	
Moved to mutual funds	505	1	0.0	
All performance-related delistings		4188	100.0	99.8

quently, information about liquidation was incorporated into the stock price before delisting, and we do not expect the missing returns to be significantly biased.

Performance-related delistings are much more frequent among Nasdaq stocks than among NYSE and AMEX stocks. There are, on average, 220 performance-related delistings from Nasdaq each year. Shumway (1997) finds only 32 such delistings each year from the NYSE and AMEX combined. This

Table II

Delistings by Year

Shown are the number of performance-related delistings on Nasdaq each year from 1973 to 1995. Data are from CRSP's 1995 daily Nasdaq file. The number of performance delistings includes 869 temporary delistings, for which CRSP does not calculate a return.

Year	Number of Delistings	Number of Firms Listed	Percentage Delisted
1973	512	3211	15.9
1974	208	2800	7.4
1975	82	2691	3.0
1976	53	2742	1.9
1977	82	2763	2.9
1978	69	2758	2.5
1979	54	2767	1.9
1980	101	2986	3.3
1981	83	3392	2.4
1982	179	3462	5.1
1983	174	4066	4.2
1984	252	4328	5.8
1985	333	4455	7.4
1986	366	4864	7.5
1987	274	5025	5.4
1988	339	4910	6.9
1989	329	4641	7.0
1990	332	4465	7.4
1991	330	4396	7.5
1992	367	4490	8.1
1993	159	4749	3.3
1994	172	5080	3.3
1995	204	5361	3.8
Average	220	3931	5.6

translates to 1.2 percent of NYSE and AMEX stocks per year on average as opposed to 5.6 percent of Nasdaq stocks, so we expect the Nasdaq delisting bias to be about 4.7 times larger than the NYSE and AMEX bias.

The fraction of Nasdaq stocks delisted each year for performance reasons does not appear to be decreasing over time. Table II lists the number of performance-related delists, the number of firms, and the ratio of delists to listed firms each year. Most of the missing returns occur in the second half of the 1980s and the early 1990s. There are also an extraordinary number in 1973. This does not appear to be a case of CRSP miscoding since the number of delists for other reasons (e.g., merger, liquidation) in that year does not appear abnormal.

B. Are Delistings Anticipated?

If impending delistings are announced while the stock is still trading, then the missing returns may not introduce a bias. At the announcement, the effect of the impending delisting will be incorporated into the price of the

stock, and we can expect the unobserved return on the delisting date to be unimportant. It is, therefore, necessary to determine how frequently investors know about impending delistings. To do this, we search Bloomberg's company news database for articles about the 159 firms that were delisted in 1993 for performance-related reasons. Only two articles indicate any predelisting announcement; 41 stories actually indicate that the delisting is a surprise. For 81 firms, we find stories about the company, but no mention of delisting; for 28 we find no stories at all; and for seven we find stories about the impending merger or liquidation (which implies an impending delisting). Therefore, it appears that most delisting events are unanticipated.

C. Temporary Delistings

In many cases, stocks that are delisted from Nasdaq are subsequently relisted on Nasdaq. In these cases, CRSP gives the stocks a delisting code that indicates that the stocks are actively traded (after the stocks are relisted). When a stock ceases trading for 10 or fewer trading periods, CRSP usually calculates a return for the period based on the price before and after the nontrading interval. When the nontrading interval is more than 10 periods, CRSP's stated policy is to ignore any change in price over the interval, and the return for the nontrading interval is coded as missing. In many cases, CRSP does not calculate a return even when the nontrading interval is less than 10 periods.[3] We find 869 cases of temporary delistings with missing returns. Some of the nontrading intervals are quite long: The longest is more than 20 years, and 595 are longer than 10 months.

Dealing with temporary delistings is somewhat problematic. One cannot simply calculate a return using CRSP price data before and after the nontrading interval; CRSP could have done so and presumably chose not to for good reasons.[4] Similarly, we do not attempt to determine a return from other sources since the period is usually too short to collect accurate data. The task is further complicated because these delistings are revealed to be temporary only after the firm is relisted. Therefore, to treat temporary delistings

[3] Consider, for example, the case of Lexicon Corporation of Fort Lauderdale, Florida (CRSP permno = 49286). On April 8, 1993, Lexicon announced that it had been delisted by Nasdaq because it failed to meet their requirements for continued listing. Lexicon had been granted an exemption to Nasdaq's listing requirements, but its exemption expired on March 31. Its stock price was $2.00 at the end of March. On August 2, Nasdaq's hearing-review committee decided to allow Lexicon's stock to be relisted. By the end of August the stock was trading at $1.625. Though Lexicon's stock lost 19 percent of its value over the five-month period that it was not listed, CRSP does not record a return for the interval.

[4] For instance, it is not clear that we are dealing with the same security or even the same firm before and after the temporary delisting. In at least some cases, the firm is reorganized during the delisting, the old equity is declared worthless, and a new equity issued. It is also highly likely that temporarily delisted firms execute reverse stock splits which make the pre- and postdelisting prices incomparable.

differently than permanent delistings would be to condition on a future event that is not known with certainty at the time of delisting, and would impose an unacceptable look-ahead bias.

Consequently, we conclude that the best way to deal with temporary delistings is to treat them in the same way we treat permanent delistings. We include temporary delistings in our sample of delistings even though CRSP does not explicitly code them as delistings. Because temporary delistings are not given a delisting code, we are unable to classify them precisely. It seems most likely that the majority occur for performance-related reasons, so we classify them as performance-related delistings. Including temporary delistings yields a total of 5,048 missing returns for performance-related delistings.[5]

II. Correcting for the Delisting Bias

A. Delisting Prices from the Pink Sheets and Other Sources

We collect over-the-counter price data for 3,330 stocks permanently delisted between 1977 and June of 1994 from semiannual summaries of the *Pink Sheets*. This sample period is shorter than CRSP's Nasdaq sample period. It begins in 1977 because *Pink Sheets* data are not available to us before that time, and it ends in June 1994 to ensure that at least four issues of the semiannual *Pink Sheets* are available following delisting.

The *Pink Sheets*, published by the National Quotations Bureau, give bid and ask quotes supplied by broker/dealers that maintain inventories in the stocks in which they specialize. Because there is no mechanism to force dealers to honor the quotes they post in the *Pink Sheets*, some quotes may not reflect the prices at which dealers are willing to trade. But although the data are not completely reliable, there is no alternative source for valuing stocks that no longer trade on organized exchanges.

The *Pink Sheets* sometimes list more than one bid and ask price for a security when more than one dealer is willing to trade the security. We calculate the price in this case as an average of the highest bid, the lowest bid, the highest ask, and the lowest ask prices. Prices for delisted stocks are usually not available immediately after delisting. We use the first price found if it is within two years of the delisting. Where price data are not available within two years, we drop the observation. Where the calculated delisting return appears to be an outlier, we examine the observation more closely, using other sources such as *Bloomberg*, and the *Directory of Obsolete Securities* (1993) published by the Financial Stock Guide Service. We find many instances where firms executed reverse splits to boost their stock price so they could be relisted on Nasdaq. Whenever we identify a split, we adjust

[5] The 869 temporary delistings plus the 4,179 permanent delistings with missing returns equal 5,048 delistings.

Table III

New Delisting Returns

Price data for stocks whose delisting returns are missing from CRSP's Nasdaq data base are collected from the *Pink Sheets*, *Bloomberg*, and the *Directory of Obsolete Securities*. This table summarizes the delisting returns calculated using those price data. The first column indicates the CRSP code for the delisting, the second indicates the number of stocks in that group that have missing delisting returns, the third gives the number of stocks for which we were able to identify a new delisting return (DRET), the fourth and fifth columns show the mean and median of the new delisting returns, and the sixth column gives the *t*-statistic for the null hypothesis that the mean new delisting return is equal to zero. The sample period is 1977 through June of 1994. The number of delistings differs from that in Table II because of the difference in sample periods.

CRSP Code	Number of Stocks	DRET	Median	Mean	*t*-Statistic
500	550	391	−18.8	−20.8	−9.99
560	914	606	−37.5	−28.9	−11.28
580	600	344	−36.0	−32.5	−10.40
550	464	255	−36.0	−30.3	−6.52
561	207	134	−30.2	−26.5	−4.08
552	101	83	−23.1	−0.2	−0.01
570	139	82	−15.0	−20.8	−3.47
551	88	53	0.4	−2.6	−0.76
581	69	30	0.0	−9.1	−1.20
520	59	47	−28.0	−35.6	−7.32
510	32	20	−45.3	−45.0	−9.65
574	23	13	−40.0	−39.8	−4.01
572	21	4	0.3	−23.5	−0.94
582	17	16	−30.3	−37.1	−3.80
573	14	9	0.0	2.2	0.96
516	5	4	−26.3	−13.8	−0.35
700	3	2	−46.8	−46.8	—
584	6	3	−54.2	−66.1	−3.83
575	10	2	4.1	4.1	—
583	5	5	−72.6	−60.9	−7.67
517	3	3	−5.3	−18.1	−0.43
All codes	3330	2107	−30.0	−26.3	−20.87

the subsequent prices. Insofar as we may not have detected all such splits, some of our recorded prices and returns may be too high, which will understate the magnitude of the bias. We also check the *Directory of Obsolete Securities* for all stocks that do not have price quotes in the *Pink Sheets*. We identify 201 securities that became worthless within five years of delisting and we assign a return of −100 percent to them.

In total, we collect returns for 63 percent of delisted stocks (2,107 out of 3,330). Table III summarizes the returns. The mean returns are fairly uniform across the most populous categories, with the average across all categories being −26.3 percent, and the overall median being −30.0 percent. The distribution of delisting returns has both a mass point at −100 percent

and a fat upper tail. Mean returns for the five most common categories are all significantly different from zero at a high confidence level, and almost all of the other estimated means are negative – many significantly so. No estimated mean is significantly positive.

B. Selection Bias in the Sample of Identified Returns

Our sources yield data on 63 percent of the stocks with missing delisting returns. To estimate the average delisting return, we must make an assumption about the 37 percent for which we find no prices. We could assume that the unlocated returns are the same as the ones we find and assign an average return of −26.3 percent to all of them. However, the delisting prices we find are not a random sample; they are more likely to be for firms that suffered the least serious performance deterioration, and the 37 percent we do not find are almost certainly from firms that performed more poorly on average. The fact that we find no trace of them suggests that many became worthless. Alternatively, we could assume that all firms we could not find became completely worthless, but this is almost certainly too negative.

Rather than make either of these extreme assumptions, it seems reasonable to assume that half of the stocks for which we cannot find data are similar to the stocks for which we find data, and the other half become worthless. The first group is then assigned an average return of −26.5 percent.[6] The second group is assigned a return of −100 percent. Incorporating this, we get an average delisting return of

$$0.63(-26.3) + 0.37(0.5(-100.0) + 0.5(-26.5)) = -40 \text{ percent.} \tag{1}$$

Later, we perform a sensitivity analysis for this assumption.

C. Liquidity Effects of Delisting

Delisting from Nasdaq has a dramatic impact on liquidity if the stock is not subsequently listed on the NYSE or AMEX, which means that an investor must sell the stock at a lower price after the delisting than before. To account for this, we examine the increase in the bid-ask spread by looking at the 320 stocks for which bid-ask data are available in the *Pink Sheets* and for which bid-ask data are also available in CRSP for at least one of the last three days that the stock was listed. In the rare instances in which more than one set of bid and ask quotes is given in the *Pink Sheets*, we use the average bid-ask spread. Defining the relative bid-ask spread as

$$\frac{\text{ask} - \text{bid}}{(\text{ask} + \text{bid})/2}, \tag{2}$$

[6] This figure is calculated by averaging the category-specific mean returns in Table III weighted by the frequency with which the unfound returns fall into each category.

we find that the mean relative bid-ask spread increases from 0.41 to 0.82 when a stock is delisted for performance reasons. The median relative bid-ask spread increases by an even larger amount, from 0.33 to 0.86.[7]

D. Calculation of a Corrected Delisting Return

To avoid a delisting bias in empirical work using CRSP's Nasdaq data, a researcher must account for all the factors discussed above. We now compute a corrected delisting return that can be used to do this, and we call this return R_{Corr}. Consider a stock listed on Nasdaq at price P that can be sold at the bid price of $P(1 - 0.41/2) = 0.795P$. After delisting, the stock's price drops 40 percent to $P(1 - 0.40)$, and it can then be sold at the bid price of $P(1 - 0.40)(1 - 0.82/2) = 0.354P$.[8] The corrected delisting return—that is, the total change in the wealth of an investor holding the stock through delisting—is, therefore,[9]

$$R_{Corr} = \frac{0.354P - 0.795P}{0.795P} = -55 \text{ percent.} \tag{3}$$

Researchers can correct for the delisting bias by using this return whenever a performance-related delisting return is missing from CRSP's Nasdaq database.

III. Implications for the Size Effect

Banz (1981) and Reinganum (1981) are the first to document that stocks with smaller market capitalizations have higher average returns and that the effect is concentrated primarily in the smallest stocks. Since Nasdaq stocks have smaller market capitalizations on average than other listed stocks, they are of particular interest in understanding the size effect, and previous work has found a size effect among Nasdaq stocks. It is reasonable to sus-

[7] The bid-ask spread calculations performed here use quoted spreads rather than effective spreads. However, since Nasdaq quotes are firm whereas *Pink Sheet* quotes are not, it seems most likely that the change in effective spreads is, if anything, larger than the change in the quoted spreads we calculate here.

[8] At first blush, it may appear that this analysis is questionable because it assumes that stocks with −100 percent returns have relative bid-ask spreads of 0.82 following delisting. However, the relative bid-ask spread after delisting is actually irrelevant for these stocks—after delisting, both the bid and the ask prices are zero regardless of the assumed bid-ask spread and the relative bid-ask spread is undefined. Consequently, this calculation actually makes no assumption at all about the postdelisting relative bid-ask spread of worthless stocks.

[9] To double-check our method of including bid-ask spread effects, we also perform the calculations in another way. Using the subset of firms for which we have both predelisting and postdelisting bid-ask data, we calculate bid-to-bid returns that incorporate the bid-ask spread change directly into the analysis. The mean bid-to-bid return is −41 percent. After allowing for firms with returns of −100 percent, and after allowing for the 37 percent of firms for which we can find no data, we arrive at a corrected delisting return virtually identical to this one.

Table IV

Average Monthly Returns for 20 Size Portfolios

All Nasdaq stocks are assigned to one of 20 portfolios each year based on their market capitalization at the end of the previous year. Returns are calculated as equal-weighted averages of monthly returns for all the stocks in the portfolio. Column 2 gives the mean market capitalization of the stocks in the portfolio, and column 3 shows the frequency with which stocks are delisted. Column 4 (CRSP) shows the mean return for the portfolio calculated with CRSP data alone, column 5 (Corr.) shows mean returns using a value of −55 percent for all missing delisting and temporary delisting returns. The last two columns report *t*-statistics for a test of whether the portfolio's average return ($\bar{R}_i$) is equal to the average return of the entire sample ($\bar{\bar{R}}$). *F*-statistics and *p*-values are reported for a test of whether all portfolio average returns are equal. Data are from CRSP's 1995 daily Nasdaq file. The sample period is 1972 through 1995.

Size Portfolio	Mean Size ($mill)	Delisting Frequency (%/mo.)	Mean Return (%)		*t*-Statistic $H_0: \bar{R}_i = \bar{\bar{R}}$	
			CRSP	Corr.	CRSP	Corr.
Smallest 1	1.1	2.95	3.79	1.97	4.83	1.86
2	2.2	1.59	2.52	1.56	2.36	1.05
3	2.3	1.15	1.78	1.17	0.90	0.29
4	4.5	0.87	1.49	1.05	0.35	0.05
5	5.8	0.70	1.43	1.09	0.23	0.13
6	7.4	0.57	1.15	0.87	−0.32	−0.29
7	9.1	0.51	1.07	0.83	−0.47	−0.37
8	11.3	0.40	1.07	0.87	−0.47	−0.30
9	13.8	0.33	1.02	0.81	−0.57	−0.41
10	16.9	0.28	1.00	0.87	−0.60	−0.31
11	20.6	0.23	1.05	0.94	−0.51	−0.15
12	25.4	0.23	0.99	0.86	−0.64	−0.31
13	31.4	0.17	0.95	0.87	−0.70	−0.29
14	39.2	0.10	0.93	0.89	−0.74	−0.26
15	49.8	0.10	0.93	0.87	−0.74	−0.29
16	65.0	0.12	0.92	0.88	−0.76	−0.28
17	88.4	0.07	0.90	0.86	−0.81	−0.32
18	129.9	0.04	1.05	1.03	−0.52	0.01
19	214.5	0.02	1.10	1.09	−0.40	0.14
Largest 20	1071.2	0.05	1.09	1.06	−0.44	0.07
F					3.52	0.58
p-value					0.0001	0.930

pect that delistings are concentrated disproportionately among smaller firms, and so the delisting bias may inflate the apparent returns to portfolios of smaller stocks. In this section, we investigate whether the apparent size effect among Nasdaq stocks is sensitive to correction of the data for the delisting bias.

Table IV demonstrates that delistings are, in fact, more common for small stocks than for large ones. We assign all Nasdaq stocks to one of 20 portfolios each year based on their market capitalization at the end of the previous year. The table shows that performance-related delistings are more frequent among smaller Nasdaq stocks and that they are large and economically sig-

nificant for the smallest portfolios. On average, 2.95 percent of stocks in the smallest portfolio are delisted each month, and 1.59 percent, 1.15 percent, and 0.87 percent are delisted each month, on average, from the next three smallest portfolios.

Lamoureux and Sanger (1989) calculate average monthly returns for portfolios of Nasdaq stocks and conclude that there is a size effect among Nasdaq stocks. We repeat that work using the same methodology. Table IV displays statistics for the returns of the 20 size portfolios formed as noted above. Mean monthly returns are calculated first using the unmodified CRSP data, and then they are recalculated using the the corrected CRSP data where a return of −55 percent is inserted for all performance-related delistings that do not have return data in the CRSP database. Both delisting and temporary delisting returns are corrected.

As in Lamoureux and Sanger (1989), the unadjusted CRSP returns show a distinct pattern of higher returns for the smaller size portfolios. Our numbers differ from theirs somewhat because of the different sample periods. (Later, we redo the analysis for their sample period.) The smallest size portfolio has an average monthly return of 3.79 percent, which is well above the mean for all portfolios and very significant economically. Each of the five smallest portfolios gives some indication of an abnormally high average return. When the data are corrected for the delisting bias, however, the abnormal returns are dramatically reduced. The mean return for the smallest size portfolio drops to 1.97 percent (a reduction of 48 percent), and only the two smallest portfolios give any indication of a higher mean return.

To test the significance of these returns, we run a dummy regression where the dependent variable is the return of each portfolio each month concatenated with the mean return of all portfolios each month, and the independent variables are a constant and 20 dummies – one for each size portfolio. In this regression, the *t*-statistic for each dummy variable is a test of whether the mean return for that portfolio is significantly different from the mean return of all the portfolios. The *t*-statistics for the dummies are reported in the last two columns of Table IV. They indicate that the returns for the two smallest portfolios are significantly different from the overall mean when no delisting correction is used. When the delisting return is applied, however, none of the portfolio dummies is significant. At the bottom of the table, the *F*-statistic is reported for the significance of the overall regression, which simultaneously tests the significance of all the dummies. For the uncorrected CRSP data, the *F*-test is highly significant, rejecting the null hypothesis that all dummy coefficients are zero. The delisting correction, however, reverses the inference dramatically, so that there is no evidence to reject the null hypothesis that all the dummy coefficients are zero. Consequently, the delisting bias is crucial for the evidence of a size effect in Nasdaq for the years 1972 to 1995.

Fama and French (1992), Dichev (1998), and others have suggested that the size effect has diminished in recent years, which raises the possibility that the size effect may survive the delisting correction in the earlier part of the sample. To address this question, we redo the analysis for the subperiods

Table V

Sensitivity Analysis of Average Monthly Returns for 20 Size Portfolios

The analysis of Table IV is repeated with modifications. All Nasdaq stocks are assigned to one of 20 portfolios each year based on their market capitalization. Columns 2 and 3 show equal-weighted average monthly returns for the subperiod 1973–1985, and columns 4 and 5 show mean returns for 1986–1995. CRSP indicates that missing delisting returns are not corrected, and Corr. indicates that a return of −55 percent is used for all missing permanent and temporary delistings. Column 6 shows mean returns when only delisting returns (DRETs) are corrected and temporary delisting returns are not. Columns 7 and 8 use corrected delisting returns (R_{Corr} =) of −66 and −43 percent respectively. Columns 9 and 10 use annual returns (expressed as compounded monthly returns for comparison) rather than monthly returns. *F*-statistics and *p*-values are reported for the test of whether all portfolio average returns are equal. The data are from CRSP's 1995 daily Nasdaq file, 1972–1995 unless noted otherwise.

Size	1973–1985		1986–1995		DRET Only	R_{Corr} = −66%	R_{Corr} = −43%	Annual Returns	
Portfolio	CRSP	Corr.	CRSP	Corr.	Corr.	Corr.	Corr.	CRSP	Corr.
Smallest 1	2.87	1.26	4.98	2.44	2.23	1.94	2.54	2.90	1.85
2	2.54	1.50	2.49	1.17	1.72	1.57	1.88	2.05	1.54
3	2.04	1.17	1.44	0.71	1.27	1.16	1.39	1.68	1.28
4	1.72	0.97	1.20	0.68	1.14	1.06	1.23	1.55	1.30
5	1.75	1.08	1.02	0.64	1.16	1.10	1.23	1.57	1.36
6	1.38	1.13	0.85	0.54	0.94	0.89	0.99	1.36	1.20
7	1.38	1.14	0.67	0.43	0.89	0.84	0.94	1.30	1.15
8	1.42	1.23	0.62	0.40	0.92	0.88	0.96	1.33	1.22
9	1.11	0.95	0.91	0.63	0.86	0.83	0.89	1.27	1.16
10	1.33	1.19	0.58	0.45	0.91	0.88	0.93	1.27	1.19
11	1.24	0.79	0.80	0.68	0.98	0.96	1.00	1.33	1.28
12	1.17	1.03	0.75	0.64	0.91	0.89	0.92	1.37	1.33
13	1.14	1.04	0.71	0.65	0.89	0.88	0.91	1.29	1.24
14	1.16	1.12	0.64	0.59	0.90	0.89	0.91	1.27	1.25
15	1.06	0.99	0.77	0.72	0.90	0.89	0.91	1.20	1.17
16	1.07	1.03	0.73	0.68	0.90	0.89	0.90	1.16	1.14
17	0.98	0.94	0.80	0.76	0.88	0.87	0.88	1.22	1.20
18	1.08	1.06	1.01	0.98	1.03	1.03	1.04	1.27	1.26
19	1.13	1.11	1.08	1.07	1.10	1.10	1.10	1.35	1.35
Largest 20	0.93	0.89	1.29	1.27	1.06	1.06	1.07	1.18	1.16
F	0.99	0.08	3.72	0.78	0.82	0.33	0.96	1.19	0.18
p-value	0.470	1.00	0.000	0.744	0.685	0.998	0.502	0.258	1.00

1973 to 1985 and 1986 to 1995. The first corresponds to the sample period of Lamoureux and Sanger (1989) and makes the results directly comparable. Table V shows results of the size analysis for the subsamples. Against expectations, the size effect is stronger (in some ways) in the later subperiod than in the earlier one. The mean return to the smallest portfolio is 4.98 percent in the later subperiod as opposed to only 2.87 percent in the first. Mean returns for the other size portfolios are not as easily compared because the mean return for all portfolios is lower in the later subperiod, but the five smallest portfolios in the later subperiod give some indication of above-average returns. Moreover, the *F*-statistic for the test that all means

are equal is 3.72 ($p = 0.0001$) for the second period and 0.99 ($p = 0.470$) for the first, further indicating that the apparent size effect (before correcting for the delisting bias) has not diminished over time.

Table V also shows that correcting for the delisting bias essentially eliminates the evidence for the size effect in both subperiods. Again, only the two smallest size portfolios show evidence of above-average returns, and the overall F-statistic is statistically insignificant in both subperiods (0.08 in the earlier one, and 0.78 in the later one). Thus, the subperiod results reinforce the results from the full sample period. It appears that there was never a significant size effect in Nasdaq stocks. Any apparent size effect (in the subperiods as well as the full sample) can be attributed to the bias introduced into the data by the missing delisting returns.

A. Interpretation of the Size Effect

The meaning of the size effect has been debated since it was first identified, and our results throw additional light on the issue. One explanation for the size effect is that the higher returns earned by small firms are compensation for the higher riskiness of small firms. Berk (1995) argues that size is functioning as a proxy for the unobservable systematic risk of these firms. Similarly, Chan and Chen (1991) and Fama and French (1995) argue that the higher average returns of smaller firms are compensation for the higher risk of financial distress among small firms. Dichev (1998), however, finds evidence inconsistent with a distress explanation of the size effect.

Small Nasdaq stocks are ideal for examining the size effect because they are, as a group, the smallest and most distressed stocks for which we have reliable data. Presumably, the size effect should be most pronounced in this sample of stocks. Table IV offers evidence that size and the risk of distress are, in fact, highly correlated among Nasdaq stocks since the smallest size portfolios have a much higher delisting frequency than the larger size portfolios. Therefore, our finding of no size effect among Nasdaq stocks is evidence against the hypothesis that the size effect is due to the systematic pricing of the distress risk of smaller firms. This is consistent with the findings of Dichev (1998). Our results also tend to support the assertion of Lo and MacKinlay (1990) and Black (1993) that the size effect is an example of a data-snooping bias. Under that hypothesis, we should not expect to see the size effect outside of the sample in which it was first identified. Because Nasdaq data provide just such an out-of-sample test, the absence of a size effect there tends to support the data-snooping hypothesis.

B. Sensitivity Analyses

The analysis uses data from sources that are incomplete and somewhat unreliable, which necessitates a number of assumptions. Now we analyze the sensitivity of the results to these assumptions. Table V displays results of the size analysis under a series of alternative assumptions. The results for the 1973 to 1985 and 1986 to 1995 subperiods have already been discussed.

In the analysis to this point, we correct for both permanent and temporary delistings. We feel this is appropriate because the two are indistinguishable at the time of delisting, and it is only later that the investors learn that a delisting is temporary. It is possible, though, that temporary delistings differ intrinsically and therefore warrant a different corrected return. To judge the importance of this concern, Table V shows results of the size analysis when only permanent delistings are corrected and temporary delistings are assigned a return of zero percent (clearly an upwardly biased estimate of the true return). This makes only a small difference to the results. While the return on the smallest portfolio increases from 1.97 percent to 2.23 percent, the overall *F*-statistic is still completely insignificant, so the inferences are unchanged.

We assume that half of the stocks for which we cannot find postdelisting returns experience −100 percent returns and the other half have the same average return as the stocks for which we find postdelisting returns. Table V shows results when it is assumed that all stocks with missing returns become completely worthless (making $R_{Corr} = -66$ percent), and the case where all stocks with missing returns have the same return as the stocks we find (making $R_{Corr} = -43$ percent). Again, the average returns to the smallest portfolios change somewhat, but the overall inferences are unaffected.

Last, we address the possibility of a bias in measured returns due to bid-ask bounce. Blume and Stambaugh (1983) suggest that bid-ask bounce may bias analysis of the size effect since the bias will be most severe for small stocks, which are more thinly traded. They recommend using buy and hold returns of up to a year to eliminate the bias. We recalculate the size portfolio returns using annual buy and hold returns. All stocks are sorted into size portfolios each year as before, but we calculate annual returns for each stock by compounding the monthly returns. When a stock is delisted, its returns subsequent to delisting are replaced by returns for CRSP's value-weighted index of all Nasdaq, NYSE, and AMEX stocks. To facilitate comparison, we convert the annual return to an equivalent compounded monthly return and report these in Table V. The size effect is weakened in the uncorrected data, but is still discernable. After correction, the measured size effect is reduced even further to the point where it is almost unobservable. The *F*-statistic for the overall regression becomes even less significant, and only the very smallest portfolio gives any sign of having an above-average return. This suggests that what little evidence of a size effect remains after correction for the delisting bias in the previous analysis is at least partially explained by bid-ask bounce in the monthly data, which further strengthens our overall results.

C. Does the Corrected Return Vary by Firm Size?

It is reasonable to speculate that the corrected delisting return may vary by firm size. This could happen if small firms have −100 percent returns after delisting more often than large firms, or if small firms with *Pink Sheet* data have lower returns, or if small firms experience a larger increase in their bid-ask spread. Any of these can introduce a bias into the corrected analysis.

To address this concern, we first write a single equation for the corrected return. Define F to be the fraction of firms for which we are able to find postdelisting returns, define R_F to be the average return of the firms for which we find postdelisting returns, and define S_0 and S_1 to be the predelisting and postdelisting relative bid-ask spreads, respectively. R_{Corr} is then given by[10]

$$R_{Corr} = \frac{F(1+R_F)\left(1-\frac{S_1}{2}\right) + 0.5(1-F)(1+R_F)\left(1-\frac{S_1}{2}\right)}{\left(1-\frac{S_0}{2}\right)} - 1. \tag{4}$$

Simplifying and rearranging, we can write

$$R_{Corr} = 0.5(F+1)(1-R_F)\left(\frac{1-\frac{S_1}{2}}{1-\frac{S_0}{2}}\right). \tag{5}$$

The corrected return is a function of three factors: F, R_F , and $[1 - (S_1/2)]/[1 - (S_0/2)]$. If these factors do not vary by size, then the corrected return also does not vary by size.

Table VI displays information about each of the factors for the size portfolios. For the first two factors, all Nasdaq firms are grouped into 20 portfolios as before. For each, the second column shows the number of firms delisted from that portfolio over the entire time period. As expected, the smallest size portfolios have the highest number of delistings. However, this is not evidence that the corrected return varies across size portfolios because the frequency of delisting is not directly relevant. Column three gives the value of F for each portfolio, and we see that F is fairly constant across portfolios. There is significant variability for the portfolios of larger firms because of the low number of observations, but nothing to suggest a systematic pattern. Column four shows the mean value of R_F for each of the portfolios, and again we see no systematic pattern. For the last factor, we have a much smaller number of observations since a stock must have both pre-

[10] The denominator represents what an investor could have sold one dollar of stock for before delisting, taking into account the relative bid-ask spread before delisting. The numerator is the average price that one dollar of stock could be sold for after delisting. The first term in the numerator represents firms for which we find postdelisting returns (F firms are found, their delisting return is R_F, and the postdelisting relative bid-ask spread is S_1). The second term represents firms for which we do not find returns and are presumed to be like the firms for which we do find returns (recall that we assume 50 percent of the unfound firms are similar to the found firms). The other 50 percent of firms for which we find returns do not appear in the numerator since we assume that they become worthless and so their postdelisting value is zero.

Table VI

Components of the Corrected Delisting Return as a Function of Size

All Nasdaq stocks are assigned to one of 20 portfolios each year based on their market capitalization at the end of the previous year. Columns 2 and 5 report the number of performance delistings (temporary as well as permanent) for each portfolio. F is the fraction of delistings for which we were able to identify delisting returns. R_F is the delisting return calculated using postdelisting prices. Only Nasdaq stocks for which bid-ask data are available before and after delisting are included in the last two columns. $[1 - (S_1/2)]/[1 - (S_0/2)]$ is a measure of the change in the relative bid-ask spread around delisting. The sample period is January 1977 through June 1994. There are fewer observations than reported in Table III because data on firm size are not available for some stocks.

				Bid-Ask Spread	
Size Portfolio	Total Number Delisted	F	Mean R_F (%)	Number with Pre and Post Data	Mean $\frac{1 - \frac{S_1}{2}}{1 - \frac{S_0}{2}}$
1	883	0.60	−24.3	54	0.751
2	495	0.67	−23.8	36	0.790
3	366	0.68	−29.8	27	0.820
4	266	0.69	−28.1	26	0.647
5	219	0.68	−31.8	21	0.743
6	174	0.74	−25.2	15	0.725
7	157	0.66	−28.8	12	0.785
8	131	0.69	−40.3	10	0.737
9	93	0.67	−32.6	11	0.788
10	82	0.68	−31.8	8	0.739
11	63	0.79	−30.8	7	0.804
12	49	0.67	−34.5	5	0.826
13	47	0.72	−18.5	3	0.737
14	27	0.59	−32.5	2	0.614
15	28	0.61	−30.4	0	—
16	27	0.81	−26.5	3	0.776
17	19	0.42	−9.0	1	0.464
18	11	0.45	4.4	0	—
19	4	0.50	68.2	0	—
Largest 20	13	0.31	−14.9	0	—
Total	3154			241	

delisting and postdelisting bid-ask spread data to be in the sample. We give the number of firms separately in column five. The last column shows the mean of the third factor, and we see that it is essentially constant across portfolios.

In sum, Table VI shows no evidence that any of the three factors vary by size, so a single corrected return appears appropriate for all delisted firms.

IV. Conclusion

In this paper, we document that many returns in CRSP's database are missing for firms that are delisted from Nasdaq for performance reasons, and that firms delisted for reasons other than poor performance have a very low frequency of missing data. We collect data on 63 percent of the observations where postdelisting returns are missing and show that they are significantly negative on average, indicating that omitting these returns introduces a significant bias. This delisting bias is 4.7 times larger than the delisting bias in NYSE and AMEX data documented by Shumway (1997). We also document that the relative bid-ask spreads of firms delisted for performance reasons increases dramatically, which further reduces the wealth of investors holding these stocks. From alternative data sources, we estimate a corrected average delisting return of −55 percent for firms that are delisted for performance reasons. Researchers can use this figure wherever CRSP shows a missing return for firms delisted from Nasdaq for performance reasons to compensate for the delisting bias in the unmodified database. It is also a reasonable estimate of the return to temporarily delisted stocks that do not have delisting information in the CRSP data.

The delisting bias in Nasdaq stocks is critical to the evidence of a size effect in Nasdaq stocks. Previous evidence that smaller Nasdaq stocks earn higher returns essentially disappears when the data are corrected for missing delisting returns. We also find that the apparent size effect (when the data are uncorrected for the delisting bias) in Nasdaq stocks has not diminished over time, and that it does not survive correction for the delisting bias in the earlier or the later part of the sample. Consequently, there is no evidence that there ever was a size effect among Nasdaq stocks. Banz (1981) finds the size effect to be most pronounced among the smallest firms. Therefore, finding no size effect among Nasdaq stocks that are the smallest in the marketplace is particularly strong evidence against the economic importance of the size effect and supports the contention of Lo and MacKinlay (1990) and Black (1993) that the size effect is due to data-snooping rather than economic fundamentals.

The size of the delisting bias and its reversal of previous inferences indicates that the delisting bias is an important consideration in studies using Nasdaq data. Studies that sort stocks into portfolios based on size, distress, or similar characteristics are particularly susceptible to the bias and it is, therefore, particularly important that the returns used be corrected for the bias.

REFERENCES

Baker, H. Kent, and Richard B. Edelman, 1991, Valuation implications of AMEX listings: A joint test of the liquidity-signaling hypothesis, *Quarterly Journal of Business and Economics* 30, 87–102.

Banz, Rolf W., 1981, The relationship between return and the market value of common stocks, *Journal of Financial Economics* 9, 103–126.

Berk, Jonathan, 1995, A critique of size-related anomalies, *Review of Financial Studies* 8, 275–286.

Black, Fischer, 1993, Beta and return, *Journal of Portfolio Management* 20, 8–18.

Blume, Marshall E., and Robert F. Stambaugh, 1983, Biases in computed returns: An application to the size effect, *Journal of Financial Economics* 12, 387–404.

Center for Research in Security Prices, 1996, *CRSP Stock File Guide* (University of Chicago, Chicago).

Chan, K. C., and Nai-Fu Chen, 1991, Structural and return characteristics of small and large firms, *Journal of Finance* 46, 1467–1484.

Dichev, Ilia D., 1998, Is the risk of bankruptcy a systematic risk?, *Journal of Finance* 53, 1131–1147.

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

Fama, Eugene F., and Kenneth R. French, 1995, Size and book-to-market factors in earnings and returns, *Journal of Finance* 50, 131–155.

Fama, Eugene F., Kenneth R. French, David G. Booth, and Rex Sinquefeld, 1993, Differences in the risks and returns of NYSE and NASD stocks, Working paper, University of Chicago.

Financial Stock Guide Service, 1993, *Directory of Obsolete Securities* (Financial Information, Inc., Jersey City, N.J.).

Lamoureux, Christopher G., and Gary C. Sanger, 1989, Firm size and turn-of-the-year effects in the OTC/Nasdaq market, *Journal of Finance* 44, 1219–1245.

Lo, Andrew, and Craig MacKinlay, 1990, Data-snooping biases in tests of financial asset pricing models, *Review of Financial Studies* 3, 431–468.

Loughran, Tim, 1993, NYSE vs NASDAQ returns: Market microstructure or the poor performance of initial public offerings, *Journal of Financial Economics* 33, 241–260.

Reinganum, Marc R., 1981, Misspecification of capital asset pricing: Empirical anomalies based on earnings yield and market values, *Journal of Financial Economics* 9, 19–46.

Reinganum, Marc C., 1990, Market microstructure and asset pricing, *Journal of Financial Economics* 28, 127–147.

Shumway, Tyler G., 1997, The delisting bias in CRSP data, *Journal of Finance* 52, 327–340.