THE JOURNAL OF FINANCE • VOL. LVI, NO. 1 • FEBRUARY 2001

The Long-Run Stock Returns Following Bond Ratings Changes

ILIA D. DICHEV and JOSEPH D. PIOTROSKI*

ABSTRACT

Using essentially all Moody's bond ratings changes between 1970 and 1997, we find no reliable abnormal returns following upgrades. However, we find negative abnormal returns on the magnitude of 10 to 14 percent in the first year following downgrades. Additional results reveal that this underperformance is especially pronounced for small, low-credit-quality firms. Also, downgrades underperform in nearly all years in the sample, and a large part of the abnormal returns occur at subsequent earnings announcements. Thus, the evidence suggests that the poor returns result from an underreaction to the announcement of downgrades, rather than from lower systematic risk.

A NUMBER OF RECENT STUDIES suggest that stock price reactions to various events are often incomplete. For example, Loughran and Ritter (1995) and Spiess and Affleck-Graves (1995) find that equity issuers substantially underperform for three to five years following the offering. Other examples include Ikenberry, Lakonishok, and Vermaelen (1995), who document abnormally high returns after share repurchases, and Ikenberry, Rankine, and Stice (1996), who find positive abnormal returns following stock splits. There are several reasons to believe that bond ratings changes offer another powerful and interesting setting for the investigation of long-run returns. First, bond ratings changes are common and well-disseminated information events. Second, existing research suggests that bond ratings changes capture economically significant shifts in the firms' economic conditions. For example, the evidence in Holthausen and Leftwich (1986) suggests that the difference between the one-year *pre*-announcement returns for upgrades and downgrades is on the magnitude of 20 to 30 percent. Such return differences suggest that even a modest bias in the market's reaction to ratings changes could lead to considerable post-announcement returns.

* Dichev is from the University of Michigan Business School, and Piotroski is from the University of Chicago. We appreciate the comments of Bob Holthausen, Jake Thomas, René Stulz, two anonymous referees, and workshop participants at Carnegie-Mellon University, University of Chicago, Columbia University, Wharton, University of Michigan, University of Texas (Austin), and University of Washington. We also appreciate the financial assistance of the Paton Accounting Center at the University of Michigan Business School. We thank Jeff Doyle for his able research assistance.

Our interest is also motivated by the fact that existing research offers only sporadic and somewhat contradictory evidence on this issue. Pinches and Singleton (1978) find no reliable abnormal returns in the year after the announcement of ratings changes, although there is some weak evidence of a return reversal following the announcement. Glascock, Davidson, and Henderson (1987) find a statistically significant return reversal on the magnitude of four to five percent in the quarter following the announcement. Holthausen and Leftwich (1986) find no abnormal returns after the announcement of upgrades. However, Holthausen and Leftwich find evidence of abnormally low returns in the quarter following downgrades, indicating a drift rather than a reversal of pre-announcement returns.

The interpretation of existing results is further complicated by three factors. First, the investigation of long-run post-announcement returns is mostly a side issue for existing studies. Second, due to data-gathering considerations, the samples of existing studies are limited to larger and better-followed NYSE and AMEX firms. For example, the sample in Pinches and Singleton consists of about 200 Moody's bond ratings changes, Glascock, Davidson, and Henderson has 162 changes, whereas the sample in Holthausen and Leftwich combines Moody's and Standard and Poor's changes for a total of about 1,100 observations. These sample limitations are likely important because most long-run abnormal returns occur for smaller and underfollowed firms (e.g., Fama (1998)), where there is a greater potential for informational inefficiencies. Third, a number of recent studies find that the delayed reactions in returns persist over long horizons, for up to several years after the announcement (e.g., Loughran and Ritter (1995)). In contrast, existing evidence is mostly limited to returns from the first quarter after the bond ratings changes.

In this study, we provide a comprehensive investigation of the long-run returns following bond ratings changes. We use a sample that comprises essentially all available Moody's bond ratings changes during 1970 to 1997. Our full sample comprises about 4,700 observations, including many small, Nasdaq-listed, and lower credit quality firms where analyst and investor following is expected to be low. The inclusion of these more marginal firms suggests that potential inefficiencies are likely to be more pronounced in our sample. Finally, the nature of our sample implies that our results are unlikely to be changed due to sample considerations or additional data-gathering efforts.

We examine three-month, six-month, first-year, second-year, and third-year abnormal stock returns following bond ratings changes. Abnormal returns are calculated as both cumulative abnormal returns (CARs) and buy-and-hold returns (BHARs), after controlling for size and book-to-market. We find no reliable abnormal returns for stocks with upgrades. However, we find substantial negative abnormal returns following downgrades. Most of the underperformance of downgrades occurs in the first year after the announcement, with negative abnormal returns on the magnitude of −10 to −14 percent a year. The underperformance of downgrades seems to extend to the second and third year after the announcement, but the abnormal returns are substantially tempered to −4 to −6 percent annually, and their significance is sensitive to alternative specifications. Further evidence reveals that the underperformance of downgrades is more pronounced for small firms and firms with noninvestment grade debt.

We also probe deeper into whether the abnormal returns following downgrades are a compensation for risk or due to some other explanation. Lacking a well-defined and accepted model of risk, it is difficult to reject a risk-based explanation. However, the documented pattern of results suggests that stock prices possibly underreact to the information content of downgrades because the poor returns are of limited duration, and are most pronounced for small and low-credit-quality firms. The results from additional tests also support this explanation. First, we find that current bond ratings changes predict changes in future fundamentals, specifically future earnings changes and bond ratings changes. More importantly, we document predictable negative returns at the subsequent earnings announcements of downgraded firms. Thus, it seems that the market does not fully anticipate the negative implications of downgrades for future profitability. Second, we find poor returns following downgrades in nearly all years in our sample, which further questions a systematic risk explanation. Summarizing, the combined evidence of this study suggests that the poor returns following downgrades are most likely due to a market underreaction to downgrade announcements.

The remainder of the paper is organized as follows. Section I describes the data. Section II describes the methodology and presents the main results for the long-run returns following bond ratings changes. Section III examines the long-run returns conditional on size, credit quality, and recent earnings changes. Section IV provides additional evidence about the nature and the causes of the abnormal post-announcement returns. Section V concludes.

I. Data and Descriptive Statistics for the Sample

In this study, we use data gathered from Moody's Default Risk Service, which is a comprehensive computerized database containing all Moody's bond ratings changes between 1970 and 1997.[1] Figure 1 illustrates the derivation of our sample. Moody's issued a total of 49,339 bond rating changes for U.S. firms from 1970 to 1997. However, many of these changes pertain to differ-

[1] The two major bond rating agencies in the United States are Moody's and Standard and Poor's. We do not include Standard and Poor's bond ratings in our sample for two reasons. First, the corresponding product offered by Standard and Poor's only covers firms whose bonds are still being rated. In other words, the Standard and Poor's database suffers from a survivorship bias, which would be problematic in our investigation of long-run stock returns. Second, Moody's and Standard and Poor's ratings are good substitutes for each other (e.g., Holthausen and Leftwich (1986) and Jewell and Livingston (1998)). This evidence suggests that changes in Moody's and Standard and Poor's debt ratings are likely to be closely related as well. Because Moody's and Standard and Poor's ratings changes do not coincide in time, there is likely some advantage from using data from both sources when examining the stock returns at the announcement of bond ratings changes. However, this advantage is substantially eroded for an examination of long-run returns.

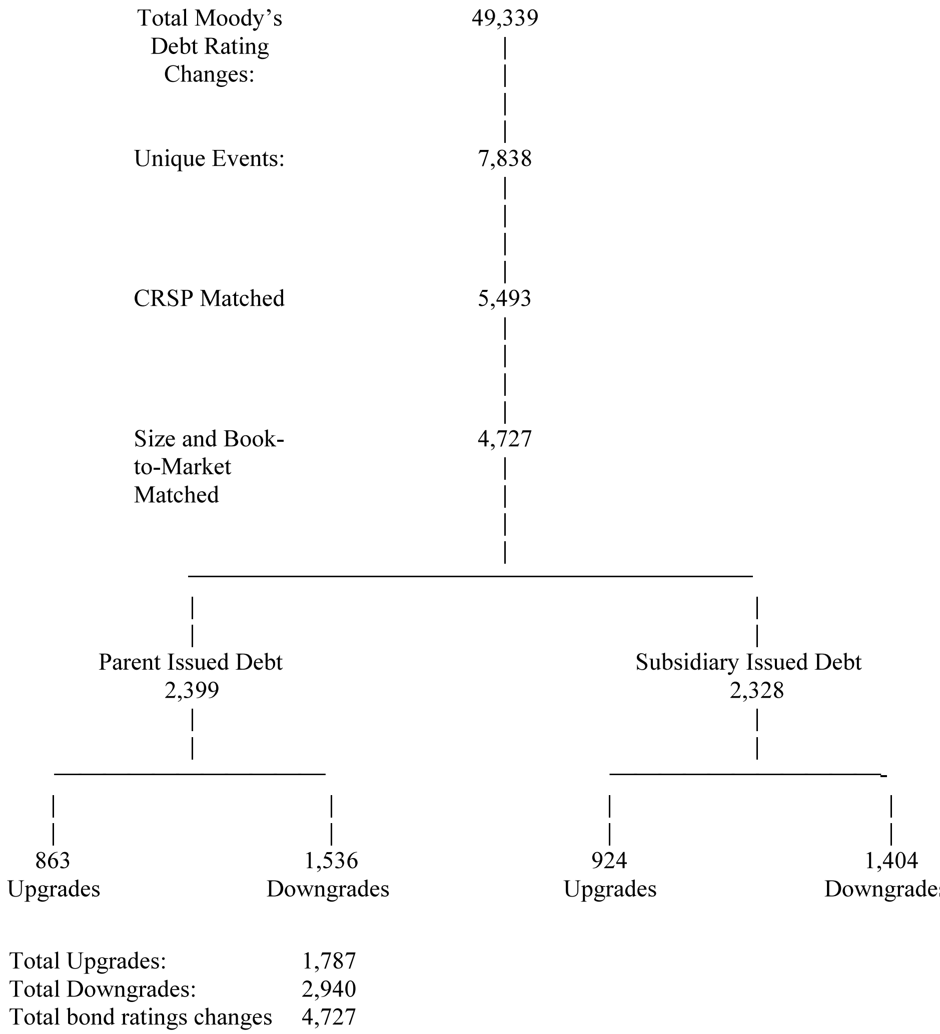


Figure 1. Derivation of the sample of Moody's bond rating changes for 1970 to 1997.

ent bond issues for the same issuing firm on the same date. In the case of simultaneous rating changes for the same firm, we retain the rating change for the most senior issue available. Eliminating such redundancies yields 7,838 unique bond rating change events. Matching Moody's data to CRSP returns yields a sample of 5,493 bond ratings changes with available long-run returns data.[2]

[2] The majority of the 2,345 bond ratings changes for which we could not find matching CRSP returns occur for private firms. Specifically, 1,141 of these bond ratings changes relate to firms that have CRSP returns for earlier or later dates but not during the event period in our study.

Our last sample requirement is motivated by several recent studies that suggest that it is important to control for size and book-to-market in computing abnormal long-run returns (e.g., Fama and French (1992)). The size requirement is entirely satisfied with CRSP data and does not lead to an appreciable loss of observations. However, the book-to-market requirement is more problematic. A simple match with COMPUSTAT book values leads to a great loss of observations. This loss is especially undesirable for our study because the firms that lack COMPUSTAT coverage are likely to be small, underfollowed, and otherwise marginal firms. Ex ante, one would expect that possible market imperfections and abnormal long-run returns are more pronounced for such firms. Thus, to avoid the loss of these observations, we complement the COMPUSTAT data with hand-collected book value data from various Moody's Manuals.[3] This process yields a final sample of 4,727 bond ratings changes with matching CRSP returns and size and book-to-market data.

The matching process also reveals important differences in debt issue characteristics, prompting us to organize the debt issues into two distinct subsamples: parent-issued debt and subsidiary-issued debt. As shown in Figure 1, the matching of parent-issued debt observations with CRSP returns and size and book-to-market data yields 2,399 observations, split between 863 upgrades and 1,536 downgrades. The matching of subsidiary-issued debt with CRSP returns is more complicated. When the subsidiary that issued the debt has publicly traded equity, we use the subsidiary's stock returns. This approach yields 1,111 observations. For those subsidiaries with no traded equity, we use the returns of the ultimate parent company, when available. This procedure accounts for another 1,217 observations. In total, we have 2,328 usable subsidiary debt observations, split between 924 upgrades and 1,404 downgrades.[4]

We distinguish between these two subsamples because it seems plausible that the price effects of bond rating changes are different, and likely weaker, for subsidiary debt issues. For example, because parent firms might provide explicit or implicit guarantees for a subsidiary's debt, it is possible that the market's reaction to a subsidiary downgrade is less severe than the reaction to a parent downgrade. Also, it is unclear whether and to what extent subsidiary rating changes translate into credit revisions and stock price effects for parent firms. Thus, the differences between parent- and subsidiary-

Apparently, these bond ratings changes occur before these companies go public (or CRSP initiates coverage) or after they are taken private. Another 1,169 of the bond ratings changes relate to firms that do not have publicly traded common equity. The remaining 35 changes relate to firms that are traded on exchanges other than the NYSE, AMEX, and Nasdaq.

[3] Specifically, we used Moody's Industrial Manual, Moody's Bank and Finance Manual, Moody's Transportation Manual, and Moody's Public Utility Manual. Moody's Manuals provided book value data on 480 additional observations.

[4] Consistent with Holthausen and Leftwich (1986) and Blume, Lim, and MacKinlay (1998), downgrades substantially outnumber upgrades in our sample. Blume et al. (1998) investigate the preponderance of downgrades, and conclude that it is at least partly due to more stringent recent standards.

Table I

Descriptive Statistics for the Sample of Moody's Bond Ratings Changes

Table I presents descriptive statistics for the full sample of bond ratings changes, spanning 1970 to 1997. Panel A presents the size and book-to-market characteristics. Market value of equity (MVE) equals total number of shares outstanding times beginning of the fiscal year stock price per share, and is presented in millions. Book-to-market (BM) equals the book value of common equity at the beginning of the fiscal year containing the bond rating change, scaled by MVE. Book values are from COMPUSTAT and Moody's Manuals. Firms with negative book value are eliminated from the descriptive statistics. Panel B presents a breakdown of the sample by exchange listing.

Panel A: Size and Book-to-Market Characteristics for All CRSP Firms, Upgrade Firms, and Downgrade Firms

	Mean	Standard Deviation	5%	25%	50%	75%	95%
All CRSP Firms							
MVE	7,743	11,572	3	14	59	332	6,040
BM	3.89	61.45	0.005	0.34	0.71	1.33	3.46
Upgrade Firms							
MVE	4,065	8,347	75	424	1,343	3,614	16,400
BM	0.69	0.47	0.18	0.39	0.63	0.90	1.39
Downgrade Firms							
MVE	2,598	5,545	19	141	631	2,403	12,353
BM	1.12	1.30	0.28	0.59	0.89	1.30	2.64

Panel B: Exchange Listing

	Upgrade Firms	Downgrade Firms	Total	Percent of Total
NYSE	1,517	2,389	3,906	82.6%
AMEX	60	198	258	5.5%
NASDAQ	210	353	563	11.9%
Total	1,787	2,940	4,727	100%

issued debt motivate separate tests for these two subsamples later in the paper. To our knowledge, this is the first study to investigate the importance of the parent/subsidiary distinction for the stock price effects of bond ratings changes.

Because we investigate long-run returns, it is important to consider the effect of known predictors of returns, most notably firm size and book-to-market (Fama and French (1992)). Panel A in Table I presents descriptive statistics for size and book-to-market for the benchmark sample of CRSP firms, as well as for the upgrade and downgrade firms in our full sample. Panel A reveals that upgrade firms are considerably larger than downgrade firms. Considering the means and the percentiles of the distribution of firm size, on average upgrade firms are about twice as large as downgrade firms. Panel A also reveals differences in book-to-market. Upgrade firms clearly have lower book-to-market than downgrades. Comparisons with the book-to-market of CRSP firms are more difficult, because CRSP firms have a much larger dispersion in book-to-market than either upgrade or downgrade firms. Overall, the differences documented in Panel A suggest that one needs to control for size and book-to-market when examining the relative performance of upgrade and downgrade firms.

The percentiles of the empirical distribution of size in Panel A of Table I reveal that our sample contains firms with a wide variation of market capitalization, including a number of small firms. However, note that most upgrades and downgrades are above the 25th percentile of size on CRSP. Thus, our sample firms are fairly sizable, which implies that the calculated stock returns later are less prone to problems of interpretation related to various types of transaction costs (e.g., judging the economic significance of abnormal returns is often problematic for very small firms because of relatively large trading costs). The impression that most firms with bond ratings changes likely do not face extreme information and transaction costs is confirmed by the exchange listing data in Panel B of Table I. Note that about 80 percent of stocks with bond ratings changes are listed on NYSE.

To further illustrate the properties of our sample, Table II presents the distribution of bond ratings changes over time for our full sample. The distribution of bond ratings changes for the parent firm sample is similar, so it is not presented. Table II reveals two major time-series properties of the data. First, there are large differences in the number of bond ratings changes across years. The number of available observations is low during the 1970s, but it rapidly increases during the early 1980s, and remains high and fairly variable during the late 1980s and the 1990s. Second, there are considerable differences in the relative proportions of upgrades and downgrades across years. The proportion of upgrades to downgrades in the full sample is about 2:3. However, as one might expect, this ratio becomes more skewed to upgrades during economic expansions (e.g., upgrades outnumber downgrades after 1992) and more skewed to downgrades during recessions (e.g., 1990 and 1991).

Table III presents a transition matrix for the full sample of bond ratings changes. The transition matrix for the parent firm sample is similar, and is thus omitted. Rows of the matrix signify the old bond rating, columns of the matrix signify the new bond rating, and the number in each cell represents the number of observations that have the respective old and new bond rating. An examination of Table III reveals the following highlights. The main diagonal of the matrix, which contains within-the-same-class ratings changes (e.g., from A2 to A3), accounts for 2,050 observations, or about 43 percent of the sample.[5] The vast majority of across-class ratings changes (e.g., from A3 to Baa1) occur within one class (within-one-class rating changes are on the two diagonals immediately above and below the main diagonal). Only 150 observations, or about 3 percent of the sample, are across two or more rat-

[5] Within-class rating changes have only been used by Moody's since April 20, 1982.

Table II
The Distribution of Moody's Bond Ratings Changes Over Time (Full Sample)

Year	Number of Upgrades	Number of Downgrades	Number of Bond Ratings Changes	Percentage of Total Bond Ratings Changes
1970	9	12	21	0.44
1971	11	18	29	0.61
1972	16	7	23	0.49
1973	13	13	26	0.55
1974	22	40	62	1.31
1975	32	26	58	1.23
1976	18	14	32	0.68
1977	23	27	50	1.06
1978	19	10	29	0.61
1979	20	34	54	1.14
1980	14	46	60	1.27
1981	21	53	74	1.57
1982	36	144	180	3.81
1983	75	99	174	3.68
1984	105	132	237	5.01
1985	100	153	253	5.35
1986	93	247	340	7.19
1987	84	185	269	5.69
1988	91	218	309	6.54
1989	90	216	306	6.47
1990	65	322	387	8.19
1991	73	236	309	6.54
1992	94	164	258	5.46
1993	128	106	234	4.95
1994	113	100	213	4.51
1995	131	113	244	5.16
1996	173	95	268	5.67
1997	118	110	228	4.82
Total	1,787	2,940	4,727	100%

ings classes. Additionally, about 40 percent of the new ratings for our sample are below Baa, which is the cutoff for investment-grade bonds. Summarizing, the descriptive evidence in Tables I, II, and III suggests that our sample has a substantial variation in firm size, magnitude of ratings changes, and credit quality. These characteristics are likely related to the strength of the information effects of bond rating changes, which implies that this sample potentially allows for powerful tests of these effects.

Previous studies examining the information content of bond ratings changes use hand-collected samples, which are necessarily somewhat limited in size and coverage. For example, Weinstein (1977) and Pinches and Singleton (1978) utilize samples of 100 and 207 Moody's bond rating change observations,

Table III

Bond Rating Change Matrix (Full Sample)

This table presents descriptive evidence on the magnitude of Moody's bond rating changes for the full sample of 4,727 observations between 1970 and 1997. Rows represent the original rating assigned by Moody's, columns represent the new rating assigned by Moody's after the change, and the number in each cell represents the number of observations that have the respective old and new bond rating. The diagonal of the matrix captures within-class rating changes. Moody's did not have the finer within-class rating designations (1, 2, and 3) before April 20, 1982, and accordingly, there are no within-class changes before that date. The proportion of downgrades to total within-class rating changes is presented in the panel below the main table.

	New Bond Rating								
	Aaa	Aa	A	Baa	Ba	B	Caa	Ca	C
Old Bond Rating									
Aaa	0	71	4	0	0	0	0	0	0
Aa	35	183	312	6	1	0	0	0	0
A	1	149	629	433	27	2	0	1	0
Baa	1	6	264	544	234	14	0	0	0
Ba	0	0	5	189	262	255	15	1	0
B	1	0	2	16	168	428	269	27	1
Caa	0	1	2	5	4	51	4	88	2
Ca	0	0	0	3	0	1	3	0	6
C	0	0	0	0	0	1	0	0	0
Proportion of "within-class" rating changes which are downgrades (%)									
	n/a	68.3	52.8	53.3	50.8	62.9	75.0	n/a	n/a

respectively. The most exhaustive studies to date (Holthausen and Leftwich (1986) and Hand, Holthausen, and Leftwich (1992)) investigate samples of approximately 1,100 observations, combining bond ratings changes from both Moody's and Standard and Poor's during 1977 to 1982. As a result of the data collection constraints, the samples of existing studies are limited to NYSE and AMEX-listed stocks, and the firms tend to be larger and have better credit quality. In comparison, Moody's Default Risk Service database yields a sample of 4,727 testable observations between 1970 and 1997, with stocks trading on NYSE, AMEX, and Nasdaq. Thus, depending on whether and how one considers the coverage of years, type of firms, number of bond ratings changes, and the redundancies from using two ratings for the same bond, we estimate that our sample is at least four to five times more comprehensive than those of previous studies.

The use of this comprehensive sample offers several advantages. First, the larger sample increases the statistical power of the tests. Second, a consideration of the relative benefits and costs of monitoring suggests that it is plausible that bond ratings changes have higher information content for smaller, Nasdaq-listed, and lower credit quality firms. Thus, the inclusion of such firms allows for more powerful tests in an economic sense. Finally, given the nature of the sample, it seems unlikely that the results of this study can be changed due to additional data-gathering efforts.[6]

II. The Long-run Returns Following Bond Ratings Changes

We present two types of evidence about the long-run returns following bond ratings changes. First, we group upgrade firms and downgrade firms into portfolios and track mean portfolio abnormal returns following ratings changes for different time horizons. This relatively simple evidence illustrates the magnitude and the duration of potential abnormal returns. However, simple statistical tests of portfolio means might be misleading due to possible cross-sectional dependencies in returns (discussed in more detail later). Therefore, we also implement two variations of Fama–MacBeth regressions to formally test the significance of abnormal returns following bond ratings changes.

Recent methodological studies disagree on the best method to calculate long-run abnormal returns. For example, Barber and Lyon (1997) favor BHARs because they reflect the compounding in long-run returns. However, Fama (1998) recommends CARs because they have better statistical properties and generally allow for cleaner tests of mispricing. On balance, it seems that both CARs and BHARs have their own strengths and can be considered as complementary rather than competing approaches to computing abnormal returns. Thus, we present both CARs and BHARs for our main portfolio results in Table IV.

Recent studies show that size and book-to-market are important determinants of the cross section of stock returns (e.g., Fama and French (1992)). Following this evidence, we calculate abnormal returns for both the CAR and the BHAR specification after controlling for size and book-to-market. Specifically, our CAR methodology is similar to the one in Brav and Gompers (1997). Each calendar month starting in January 1970, we form 25 (5×5) value-weighted portfolios of all NYSE, AMEX, and Nasdaq stocks based on their size and book-to-market. First, we divide the monthly cross sections into size quintiles. Size is measured as closing prices from the previous month times the most recent number of shares outstanding. The size quintile breakpoints are based on NYSE firms only. Within each size quintile, we form five book-to-market portfolios. Book values equal the last reported book value for

[6] The use of Moody's computerized database has one disadvantage relative to previous studies. Currently, Moody's database offers no corresponding equivalent to the sample of additions and resolutions to the Standard and Poor's Credit Watch list in Holthausen and Leftwich (1986) and Hand et al. (1992). Companies are added to the Credit Watch list when a ratings change is likely, and removed when the case is resolved. Thus, the omission of a Credit Watch list type of data might understate the true reaction to the announcements of bond ratings changes. However, the evidence in Holthausen and Leftwich (1986) reveals that most rating changes are not preceded by a Credit Watch listing, and that most entries in the list are resolved within three months. Therefore, the omission of such data is less of a problem for a long-run investigation.

Table IV

The Long-Run Abnormal Stock Returns Following Moody's Bond Rating Changes

Firm-specific BHARs are measured as the buy-and-hold raw return for the appropriate horizon minus the buy-and-hold return for a benchmark portfolio matched on size and book-to-market. Firm-specific CARs are the cumulations of monthly firm-specific returns minus the corresponding monthly return for the matching size and book-to-market portfolio. The abnormal returns presented in the table are the means of firm-specific abnormal returns for appropriate horizons. For comparative purposes, we also present the related three-day announcement period returns for these firms. t-statistics are presented in parentheses. Panel A presents results for all firms between 1970 and 1997 (N = 4,727); Panel B presents results for the parent-firm sample (N = 2,399).

	3-day Annc. Returns	3-month Returns	6-month Returns	1st-year Returns	2nd-year Returns	3rd-year Returns
Panel A: All Firms						
Upgrade firms						
BHAR	0.0048	0.0013	0.0001	0.0033	0.0101	−0.0237
(t-statistic)	(5.829)	(0.349)	(0.012)	(0.396)	(1.288)	(−3.043)
CAR	—	0.0019	0.0043	−0.0093	−0.0053	−0.0243
(t-statistic)	—	(0.546)	(0.879)	(−1.288)	(−0.719)	(−2.960)
Downgrade firms						
BHAR	−0.0197	−0.0395	−0.0662	−0.1014	−0.0368	−0.0423
(t-statistic)	(−11.47)	(−11.04)	(−7.939)	(−10.47)	(−3.842)	(−5.074)
CAR	—	−0.0402	−0.0647	−0.0957	−0.0448	−0.0339
(t-statistic)	—	(−8.108)	(−9.350)	(−10.251)	(−4.606)	(−3.744)
Panel B: Parent Firms						
Upgrade firms						
BHAR	0.0047	−0.0013	−0.0027	−0.0010	−0.0092	−0.0289
(t-statistic)	(3.643)	(−0.233)	(−0.333)	(−0.072)	(−0.762)	(−2.572)
CAR	—	0.0012	0.0024	−0.0120	−0.0272	−0.0306
(t-statistic)	—	(0.229)	(0.308)	(−1.097)	(−2.363)	(−2.500)
Downgrade firms						
BHAR	−0.0251	−0.0516	−0.0922	−0.1416	−0.0645	−0.0692
(t-statistic)	(−8.905)	(−6.924)	(−9.358)	(−9.728)	(−4.407)	(−5.521)
CAR	—	−0.0506	−0.0877	−0.1285	−0.0692	−0.0614
(t-statistic)	—	(−6.744)	(−8.010)	(−8.617)	(−4.606)	(−4.377)

a period ending at least six months prior to the ratings change. Book values are assigned in the following order: first, from the quarterly COMPUSTAT tapes, second, from the annual COMPUSTAT tapes, third, as explained earlier, from Moody's Manuals. Based on the size and the book-to-market quintile cutoffs, each month we assign all firms with no ratings changes into one of the 25 (5×5) portfolios and calculate value-weighted returns. At the end, for each month of our sample period we have 25 portfolio returns stratified by size and book-to-market characteristics.

Next, firms with bond ratings changes are assigned monthly into one cell of the 5×5 size and book-to-market matrix of benchmark portfolio returns. A monthly abnormal return equals the firm-specific return for that month minus the return on the matched size and book-to-market benchmark portfolio for that firm and month. Firm-specific returns include delisting returns (for both CARs and BHARs). Monthly firm-specific abnormal returns are added to form three-month, six-month, first-year, second-year, and third-year firm-specific CARs. The reported returns in Table IV represent the means of the firm-specific CARs for the appropriate time horizon. Table IV also reports t-statistics for whether the mean CARs are significantly different from zero.

BHARs are measured as firm-specific buy-and-hold returns minus the buy-and-hold return from the corresponding size and book-to-market benchmark portfolio (from the 5×5 matrix explained above). Return compounding starts the second day after the announcement of the bond rating change, and (similar to Loughran and Ritter (1995)) ends on the earlier of the last day of CRSP trading or the last day of the respective return window.[7] BHARs reported in Table IV are means of firm-specific BHARs for the appropriate sample and time horizon. We use simple t-statistics to test reported BHARs for significance. We do not attempt any explicit adjustment for the problems in long-run returns identified by Kothari and Warner (1997) and Barber and Lyon (1997). The reason is that Kothari and Warner (1997) and Barber and Lyon (1997) reveal that problems associated with long-run BHARs appear in the three- to five-year horizons. In our paper, none of the return windows is longer than one year. For comparison with existing studies, we also present the three-day buy-and-hold return at announcements of bond ratings changes.

An examination of the three-day mean announcement period returns in Table IV reveals that the sign and the magnitude of the announcement returns in our study are consistent with existing evidence (e.g., Holthausen and Leftwich (1986) and Hand et al. (1992)). Downgraded firms have a significantly negative mean announcement return of −1.97 percent for the all firms sample (and −2.51 percent for parent firms). Upgraded firms have a small positive mean announcement return of about 0.5 percent. The only substantial difference from existing results is that we find that the positive announcement returns for upgrade firms are significant. Because the magnitudes of the positive returns are similar for existing research and for our study, the statistical significance in our study is most likely due to our substantially larger sample.

Panel A in Table IV presents the long-run abnormal returns following upgrades and downgrades for the full sample of 4,727 rating changes between 1970 and 1997. As discussed earlier, we expect that return effects are easier

[7] For each rating change, returns are compounded over both the fraction of the announcement month following the announcement period and the faction of the last calendar month that the return compounding period spans. Because we use monthly CRSP returns to calculate the benchmark portfolio, we "spliced" monthly returns to measure the benchmark returns within these fractional months. Results and inferences are substantially the same if compounding starts on the first day of the month following the rating change.

to interpret and possibly stronger for the parent firm subsample. Thus, we present separate results for the parent firm sample in Panel B. An examination of Table IV yields two key findings. First, there are no reliable abnormal returns following upgrades. The economic magnitude of calculated abnormal returns is close to zero for both samples, using both CARs and BHARs, and over all horizons. *T*-statistics are also insignificant throughout, with the exception of those for third-year returns.

Second, bond downgrades are followed by substantial negative abnormal returns that persist for up to three years after the announcement. The underperformance of downgrade firms seems most concentrated in the first months after the announcement, gradually tapering off over the course of the first year, and further reduced over the second and third subsequent year. Specifically, Panel A of Table IV shows that for downgrade firms, both CARs and BHARs become more negative throughout the first year, and by the end of the first year the abnormal returns reach about -10 percent. The underperformance seems to extend to the second and third year after downgrades, but it is substantially tempered to about -3 to -4 percent annually.

The pattern of results for parent firm downgrades in Panel B of Table IV is consistent with that for all firms in Panel A, except that the negative abnormal returns following downgrades are more pronounced at all horizons. This evidence is consistent with our earlier conjecture that the return effects are possibly cleaner and stronger for parent-issue firms. The mean CARs and BHARs for parent-issue downgrades reach about -13 to -14 percent by the end of the first year, and then they taper off to about -6 to -7 percent in the second and third year following the announcement. The negative abnormal returns are highly significant for both samples, using both returns specifications, and at all horizons. More importantly, the negative abnormal returns following the ratings downgrades are large in economic terms.

At this point, the sheer economic magnitude of the abnormal returns is more important because the simple *t*-statistics in Table IV might be misleading due to cross-sectional dependencies in returns. There are two possible sources of cross-sectional dependence in the returns of our sample. First, subsequent ratings changes often occur within the specified time horizon of post-announcement returns. Thus, the inclusion of multiple, closely occurring, firm-specific bond ratings changes implies that at least a portion of the returns will be overlapping.[8] Second, various other commonalities in

[8] In our main tests, we do not exclude overlapping returns because our mean returns have the natural interpretation of resulting from a simple strategy of investing equal dollar amounts in every stock with a bond rating change. Note that overlapping returns might increase the weighting of any particular firm's return for some time but that does not change the intended interpretation of the results. To state in different terms, cross-sectional dependence in returns due to overlapping returns is a statistical, rather than an economic, problem. Thus, this statistical problem is probably best addressed with statistical methods. This is precisely the direction we take by implementing Fama–MacBeth regressions. In any case, the tenor of the results remains the same after eliminating overlapping returns.

returns across firms (e.g., industry-wide effects) provide another source of cross-sectional dependence.

Cross-sectional dependence in returns does not bias our estimates of mean portfolio performance. However, it could result in understated standard errors and inflated t-statistics. We use Fama–MacBeth regressions to specifically address the problem of cross-sectional dependence and, more broadly, to provide an alternative specification for estimating abnormal performance following bond ratings changes. The Fama–MacBeth regressions in our study are similar to those in Loughran and Ritter (1995). Using all CRSP firms with size and book-to-market data, we regress raw one-month returns on the log of size, log of book-to-market, and an indicator variable for upgrades or downgrades. Regressions are run cross-sectionally for each calendar month between 1970 and 1997. The upgrade (downgrade) indicator variable is equal to one if the firm's debt was upgraded (downgraded) within three months, six months, one-year, one-to-two years or two-to-three years of the current month begin investigated (depending on the return horizon being examined). The distribution of the time series of the coefficients from these monthly cross-sectional regressions is used for statistical testing. Specifically, the reported coefficient from the Fama–MacBeth regressions is equal to the average of the coefficients in the monthly cross sections. The t-statistic on the coefficient is equal to the average coefficient divided by the standard error of the time series of monthly coefficients. Essentially, Fama–MacBeth regressions use cross-sectional data only for coefficient estimation (cross-sectional dependence does not affect the unbiasedness of the coefficients) but rely on time-series independence for statistical testing. Thus, the results of Fama–MacBeth regressions are immune to problems of cross-sectional dependence.

Simple Fama–MacBeth regressions weight equally each monthly cross section in estimating the average regression coefficients. This equal weighting seems somewhat objectionable for our sample because, as shown in Table II, the number of bond ratings changes differs dramatically over time. Equal weighting overweights monthly cross sections with few observations and noisy estimates of returns, and underweights the more reliable estimates from denser monthly cross sections. Fama (1998) recognizes this problem and, as a solution, recommends weighting the monthly regression estimates by their precision (see Jaffe (1974) and Mandelker (1974) for early examples of weighting by precision).[9] Thus, for completeness, we present results from both equal-weighted and precision-weighted Fama–MacBeth regressions.

The regression results are in Table V, Panel A for the all-firms sample, and Panel B for the parent-firm sample. Table V presents a multitude of results related to our two samples (parents versus all firms), type of rating change (upgrades versus downgrades), Fama–McBeth methodology (equal versus precision weighting), and different time horizons. However, the more

[9] Precision is defined as the inverse of the standard error of the coefficients.

Table V

Fama–MacBeth Regressions of Monthly Returns on Size, Book-to-Market, and Bond Ratings Changes

This table presents the average coefficients and t-statistics (in parentheses) on UPGRADE and DOWNGRADE from the monthly estimation (January 1970 to December 1997) of the following cross-sectional regression:

$$\text{RETURN}_i = \alpha + \beta_1 \log(\text{MVE}_i) + \beta_2 \log(\text{BM}_i) + \beta_3(\text{UPGRADE}_i \text{ or DOWNGRADE}_i)$$

RETURN is raw monthly return. MVE equals total number of shares outstanding times stock price per share. Book-to-market (BM) equals the book value of common equity, scaled by MVE. The indicator variable UPGRADE (DOWNGRADE) equals one if the firm's debt was upgraded (downgraded) within the specified time horizon preceding the monthly return, zero otherwise. Equal-weighted coefficients are the average coefficients from the 336 monthly regressions, and their t-statistics are equal to the average coefficient divided by its time-series standard error. Precision-weighted coefficients are defined identically, only first the monthly cross-sectional estimates are weighted by the inverse of their cross-sectional standard error.

	3-month	6-month	1st-year	2nd-year	3rd-year
Panel A: All Firms					
Upgrade					
Equal-weighted	0.0089	0.0049	0.0028	0.0020	0.0041
	(3.972)	(2.640)	(2.193)	(1.517)	(2.756)
Precision-weighted	0.0080	0.0054	0.0036	0.0025	0.0039
	(5.851)	(5.019)	(4.355)	(2.996)	(3.992)
Downgrade					
Equal-weighted	−0.0127	−0.0098	−0.0042	−0.0024	−0.0022
	(−4.539)	(−4.257)	(−2.260)	(−1.285)	(−1.680)
Precision-weighted	−0.0157	−0.0106	−0.0059	−0.0021	−0.0026
	(−6.801)	(−6.128)	(−4.387)	(−1.611)	(−2.212)
Panel B: Parent Firms					
Upgrade					
Equal-weighted	0.0075	0.0032	0.0024	−0.0001	0.0029
	(3.179)	(1.435)	(1.475)	(−0.088)	(1.762)
Precision-weighted	0.0076	0.0046	0.0032	0.0008	0.0007
	(4.781)	(3.482)	(3.359)	(0.871)	(0.351)
Downgrade					
Equal-weighted	−0.0195	−0.0130	−0.0080	−0.0011	−0.0017
	(−5.297)	(−4.107)	(−3.283)	(−1.135)	(−1.351)
Precision-weighted	−0.0233	−0.0148	−0.0088	−0.0016	−0.0022
	(−7.806)	(−6.826)	(−5.320)	(−1.429)	(−1.886)

important patterns of results in Table V can be summarized in two main points. First, unlike portfolio-level results, regression results reveal some evidence of positive and significant abnormal returns following upgrades, especially using the precision-weighted specification and for the first year. However, the economic magnitude of these results seems fairly modest, with positive abnormal returns on the magnitude of 3 to 4 percent annually at the one-year horizon. The modest economic magnitude of these returns and the lack of agreement between portfolio and regression results advise caution before alleging abnormal performance following upgrades.

Second, the regression results confirm the underperformance of downgrades. Again, the abnormally low returns are most pronounced in the first months and quarters following downgrades, and seem to extend for about a year. Annualized abnormal returns following downgrades are on the magnitude of −5 to −7 percent at the one-year horizon for the all-firms sample, and approximately −10 to −11 percent for the parent sample. All of these returns are highly statistically significant. The regression results, however, fail to find underperformance beyond the first year following downgrades. The implied annualized returns are limited to a magnitude of −2 to −3 percent, and most coefficients are not statistically significant.

To verify that our results are robust, we also investigate three other abnormal return specifications. First, it is possible that the poor returns following downgrades might be due to the poor returns following debt and equity issues (e.g., Loughran and Ritter (1995) and Spiess and Affleck-Graves (1999)). This concern seems especially relevant here because about 70 percent of our downgrades are preceded by a debt or equity issue (within five years of the downgrade). However, we find that the abnormal return results remain essentially the same after controlling for the effect of debt and equity issues.[10] Second, we find that the results remain qualitatively the same using value-weighted instead of equal-weighted returns.[11] Third, we check for abnormal returns using market-adjusted returns instead of adjusting for size and book-to-market. The results from this specification are similar to those reported in the paper. Thus, the results seem robust to alternative specifications of abnormal returns.

Summarizing, portfolio and regression results reveal little or no evidence of abnormal returns following upgrades. However, both specifications identify an economically large and statistically significant underperformance fol-

[10] We use two different specifications to control for the effect of debt and equity issues. First, we simply eliminate all downgrades with debt and equity issues within the last five years, and recalculate the results. Second, we use matched-portfolio returns. Specifically, each of our 5×5 monthly CRSP portfolios (stratified on size and book-to-market) is further split into three subportfolios depending on whether the firm had a debt issue, an equity issue, or no debt or equity issues. Then, we match these portfolio returns with the corresponding returns following downgrades to control for the effect of debt and equity issues. After redoing the tests, we find that the tenor of the results tabulated in the paper remains the same for these two additional specifications.

[11] However, the absolute magnitude of the negative abnormal returns following downgrades is lower. To illustrate, the one-year value-weighted CAR following downgrades is −3.63 percent for the all firms sample, and −5.95 percent for the parent sample, both highly statistically significant. Tabulated results later in the paper reveal that the underperformance following downgrades is more pronounced for smaller firms, which explains the pattern of value-weighted versus equal-weighted returns.

lowing downgrades. The underperformance is most pronounced in the first several months and quarters following downgrades and extends for about a year.

III. Long-run Returns Conditional on Size, Credit Quality, and Earnings Surprises

This section extends the main results by examining the long-run returns following bond ratings changes conditional on firm size, credit quality, and the sign of the most recent quarterly earnings change.[12] To lighten the presentation, we only include results for parent issues because the preceding sections document that the long-run effects for these firms are stronger and easier to interpret. Limiting the sample to parent firms also allows for hand-collecting additional data, which we need for some of the tests later in the paper. Because the main results do not differ appreciably across CARs and BHARs, for clarity and brevity we only present results for mean one-year CARs.

A. Long-run Returns Conditional on Firm Size and Credit Quality

Existing research demonstrates that abnormal return performance and possible market imperfections are stronger for smaller firms (e.g., Bernard and Thomas (1989) and Fama (1998)). We hypothesize that if the abnormal returns observed in this study are due to an incomplete market reaction to the Moody's announcement, then the effect should be stronger for small, thinly followed firms. In addition, it seems plausible that the information effects of bond ratings changes might differ for bonds of different credit quality. A given change in bond ratings translates into a larger revision of default risk for firms with poor credit quality as compared to firms with high credit quality (e.g., Keenan, Carty, and Shtogrin (1998)). Thus, if the abnormal returns following downgrades are due to an underreaction to available information, one might expect that the return effects are stronger for firms of low credit quality. Because small firms tend to have lower bond ratings (e.g., Kaplan and Urwitz (1979)), and in our sample the correlation between firm size and credit quality is fairly high (0.57), we provide results conditional on the joint distribution of size and credit quality.

[12] Another common specification in the literature is to examine the magnitude of the return effects conditional on the magnitude of the bond rating changes. In particular, if the magnitude of the rating change is indicative of the severity of the change in financial conditions, then a change across rating classes should have substantially larger effect on returns than a within-class rating change. Consistent with the findings in Holthausen and Leftwich (1986), we find that announcement-period returns are more pronounced for an "across-class" ratings change than for a "within-class" ratings change (results not tabulated). However, we do not find substantial differences between the long-run abnormal returns for across-class and within-class bond ratings changes.

Table VI

One-year Mean CARs Following Bond Rating Changes, Conditional on Firm Size and Credit Quality (Parent-issued Debt Only)

This table presents one-year mean CARs following upgrades and downgrades, conditional on firm size (market value of equity) and credit quality (investment- or non-investment-grade). Firms are classified as small (large) if their market value of equity is less than (greater than or equal to) the median market value of equity for parent firms with a Moody's debt rating change in the same year. Bonds are classified as investment grade (noninvestment grade) if the new Moody's rating is Baa or better (Ba or lower). CARs are the cumulations of monthly firm-specific returns minus the corresponding monthly return for the matching size and book-to-market quintile from CRSP. t-statistics are in parentheses. N is the number of available observations for 1970 to 1997.

Firms	Investment Grade	Noninvestment Grade
	Panel A: Upgrade Firms	
Large	CAR = −0.0355	0.0840
	t-stat. = (−2.713)	(2.481)
	N = 419	99
Small	−0.0078	−0.0129
	(−0.335)	(−0.458)
	124	221
	Panel B: Downgrade Firms	
Large	CAR = −0.0454	−0.0601
	t-stat. = (−3.347)	(−1.323)
	N = 525	148
Small	−0.0991	−0.2086
	(−3.618)	(−7.482)
	141	722

Table VI presents a 2×2 matrix of portfolio returns conditional on firm size and credit quality, separately for ratings upgrades and downgrades. Observations are assigned into the matrix based on annual independent rankings of firm size and credit quality. Firms are classified annually as large (small) if their market value of equity is more (less) than the median market value of equity of all parent firms that had a Moody's rating change during that year. Operationally, we define debt to be of low credit quality if it is noninvestment grade (i.e., its revised debt rating is Ba or lower) and to be of high credit quality if it is investment grade. The distribution of the number of observations in Table VI illustrates the strong correlation between firm size and credit quality. For both upgrades and downgrades, the observations are clearly clustered along the diagonal, which slopes down from left to right.

Table VI reveals two important findings with respect to abnormal returns. First, we find little to no evidence of abnormal returns for the finer partitions of upgrades. Large non-investment-grade firms seem to have a positive mean abnormal return, and large investment-grade firms have a negative abnormal return. However, these results look fairly fragile and unreliable, either in terms of the economic magnitude of the abnormal returns, the number of available observations, or the magnitude of the *t*-statistics. Thus, these results should be treated with caution. Second, the underperformance following downgrades is significant in most partitions, but is visibly most pronounced for the small non-investment-grade firms. This evidence is consistent with our conjecture that the underperformance is likely to be more pronounced for more marginal and underfollowed firms. The mean first-year CAR for small non-investment-grade downgrades is a remarkable −21 percent, with a *t*-statistic of −7.48. Also note that this estimate is based on 722 observations, by far the largest subgrouping in Table VI.[13]

B. Long-run Abnormal Returns Conditional on the Preceding Quarter's Earnings Surprise

A question that arises from the preceding results is whether the poor returns following downgrades might be due to a "correlated omitted variable" problem, that is, the return pattern is the result of another corporate event, or some other variable, that is correlated with both rating changes and future returns. Probably the most serious such concern is whether bond ratings changes are merely a proxy for earnings surprises. Kaplan and Urwitz (1979) show that profitability on assets is a significant predictor of bond ratings, which implies that bond ratings changes and earnings changes are likely correlated in our sample. Additionally, it is well documented that negative earnings changes are followed by abnormally low returns (the so-called post-earnings-announcement drift, e.g., Bernard and Thomas (1989)). Thus, the poor returns following downgrades might simply be due to the effect of correlated earnings changes.[14]

We attempt to distinguish between the two drift effects by investigating the abnormal returns following ratings changes conditional on the sign of the most recent quarterly earnings change. If the abnormal returns following downgrades are due to the effects of correlated earnings changes, then the underperformance following downgrades should be present only for negative earnings changes, and should disappear if the downgrades are pre-

[13] An additional analysis to map out the characteristics of these firms reveals that they are relatively small (mean MVE of $148 million), have a fairly high book-to-market (mean of 1.28), are listed more often on Nasdaq (22 percent on Nasdaq vs. 12 percent for the full sample), and have a high probability of exchange delisting (16 percent delist within a year after the downgrade), and especially of a performance delisting (10 percent). These descriptive statistics confirm the intuition that small, non-investment-grade downgrades occur for marginal firms with poor prospects.

[14] We also calculated abnormal returns after controlling for the effect of mergers and acquisitions. After excluding all firms that had a merger or acquisition within 12 months prior to a rating change, the tenor of the results remains the same. If anything, the underperformance following downgrades is more pronounced for this specification.

Table VII

One-Year Mean CARs Following Bond Rating Changes After Controlling for the Sign of the Post-Earnings-Announcement Drift (Parent-Issued Debt Only)

Table VII is organized as a 2×2 matrix that presents one-year mean CARs following upgrades and downgrades, conditional on the sign of the most recent preceding quarterly earnings change. CARs are the cumulations of monthly firm-specific returns minus the corresponding monthly return for the matched size and book-to-market quintiles from CRSP. t-statistics are in parentheses. N is the number of available observations for 1970 to 1997. Quarterly earnings changes are defined as income before extraordinary items less income from the same quarter in the previous year. The quarterly earnings realizations and announcement dates are obtained from COMPUSTAT's Quarterly Industrial tape, Compact Disclosure, annual reports, and the *Wall Street Journal Index*. The columns and the rows of the matrix are summed up to provide a reference for the return effect of bond rating changes and earnings changes, unconditional on each other. The entries in the cells of the matrix provide evidence about the abnormal return effects of bond ratings changes and earnings changes, conditional on each other.

	Upgrade Firms	Downgrade Firms	Unconditional Earnings Surprise Effect
Positive earnings surprise	CAR = −0.0032 t-stat. = (−0.233) N = 537	−0.1038 (−5.048) 599	−0.0563 (−4.412) 1136
Negative earnings surprise	−0.0298 (−1.659) 322	−0.1412 (−6.720) 910	−0.1121 (−6.889) 1232
Unconditional bond ratings change effect	−0.0132 (−1.202) 859	−0.1264 (−8.382) 1509	

ceded by positive earnings changes. Table VII presents the results from this conditional investigation. For each rating change, we calculate the change in earnings for the most recent quarterly earnings announcement preceding the rating change. Following existing post-earnings-announcement drift research, we define quarterly earnings change as income before extraordinary items (COMPUSTAT item 18) less income from the same quarter in the previous year (e.g., Foster, Olsen, and Shevlin (1984) and Bernard and Thomas (1989, 1990)). The quarterly earnings realizations and announcement dates are obtained from COMPUSTAT's Quarterly Industrial tape. When COMPUSTAT data are unavailable, we fill in the earnings data from the following sources: Compact Disclosure for years 1989–1997, annual reports (1973–1991), and the *Wall Street Journal Index* (1970–1997). In the end, we are able to find earnings data for 2,368 out of the 2,399 parent observations.

Table VII is organized as a 2×2 matrix, where one-year mean CARs following bond ratings changes are split by upgrades/downgrades and positive/negative earnings changes. The columns and the rows of the matrix are summed up to provide a reference for the unconditional return effect of bond rating changes and earnings changes. The entries in the cells of the matrix provide evidence about the effect of the bond ratings changes and the earnings changes conditional on the other variable.

Table VII reveals two important findings. First, the unconditional post-earnings-announcement drift is weak to moderate in our sample (return difference between positive and negative earnings surprises is about 5.6 percent). In contrast, the unconditional return drift following downgrades is strong, mean CAR of about −12.6 percent. Also note that the unconditional CAR following positive earnings surprises is negative, the opposite of what one would expect if the abnormal returns were simply due to the post-earnings-announcement drift. Thus, the unconditional return evidence suggests that the post-earnings-announcement drift is unlikely to account for the drift following downgrades. Second, there are no reliable abnormal returns following upgrades for either positive or negative earnings surprises. However, there are substantial and significant negative abnormal returns of −10 and −14 percent following downgrades for positive and negative earnings surprises, respectively. Thus, downgrades underperform even after positive earnings changes, which again suggests that the downgrades drift is distinct from the post-earnings-announcement drift.[15]

IV. Evidence on the Causes of the Poor Returns Following Downgrades

The preceding sections provide evidence of abnormally poor returns following bond downgrades.[16] In this section, we probe deeper into the potential causes for this phenomenon. Generally, there are three possible explanations for unusually high or low returns. First, the returns are only seemingly abnormal, but actually represent a compensation for some underlying form

[15] Downgrade announcements are typically preceded by poor stock returns (e.g., Holthausen and Leftwich (1986)). Thus, the poor returns following downgrades are essentially a continuation of the poor returns preceding downgrades. In other words, the negative drift following downgrades is akin to momentum effects (e.g., Chan, Jegadeesh and Lakonishok (1996)). This observation raises the question of whether the downgrade drift is different from momentum effects. However, it is probably clear that momentum effects cannot literally explain the results of this study because it is the underreaction to information events (such as bond ratings changes) that causes momentum effects, rather than the other way around. In any case, the tenor of the results remains unchanged after controlling for momentum (in addition to size and book-to-market).

[16] If anything, the tabulated results probably understate the underperformance following downgrades. The reason is that the tabulated results are from the 4,700 observations with available size and book-to-market data. However, we exclude the returns of another 800 observations for which we have CRSP returns but no size or book-to-market data. The mean one-year raw return following these downgrades is about −14 percent. Thus, even after any reasonable adjustment for expected returns, it seems that the exclusion of the returns of these marginal firms likely understates the underperformance following downgrades.

of risk omitted in the research design. This explanation seems strained here, because it suggests that downgrade firms underperform because they hedge against some sort of systematic risk. Additionally, this explanation implies that the value of this hedge is highest in the first months to a year after a downgrade, and it quickly deteriorates thereafter.

Second, the detected abnormal returns seem abnormal ex post but the ex post realized returns might be a biased reflection of the ex ante expectations of investors. In other words, the poor returns following downgrades could have occurred because the sample period from 1970 to 1997 brought an unexpected deterioration in the performance following downgrades. Results of other research also suggest that this explanation is plausible. For example, evidence in Altman (1993) and Dichev (1998) reveals a large economy-wide increase in corporate distress and bankruptcies since about 1980. The evidence in Dichev (1998) also suggests that this increase in corporate distress was not fully anticipated by the market. Thus, investors could have been surprised about the negative implications of downgrades during that period (recall that most of our bond ratings changes date from after 1980). However, note that this "ex ante insufficient information" explanation becomes more problematic as the sample period increases. Because our sample period is 28 years, this explanation is possible, but not overly likely. An examination of sample subperiods could shed additional light on this explanation. If the abnormal returns are due to the insufficient-information explanation, they are more likely to be confined to some subperiod of the sample, and taper off later with investors learning from their trading experience.

Third, the poor returns following downgrades could be due to market inefficiencies. In other words, the third explanation implies that investors are fully aware of the information implications of downgrades, but stock prices do not reflect this information because of various information-processing, behavioral, or institutional biases. For example, Bernard and Thomas (1990) show that investors persistently underreact to earnings surprises, although data and models about the future implications of earnings surprises are widely available. Similar behavioral biases might be behind the underreaction to downgrade announcements as well.

In this section, we search for additional evidence that could help us distinguish among these alternative explanations. First, we investigate whether current bond ratings changes predict future business fundamentals, specifically subsequent bond ratings and earnings changes. Then, we test whether the market is surprised by these predictable future changes in firm fundamentals. Second, we investigate whether the poor returns following downgrades are persistent. If downgrades earn poor returns because they hedge against some sort of systematic risk, then the abnormal returns following downgrades should be positive in at least some time periods (for the same argument see also Lakonishok, Shleifer, and Vishny (1994)).

Table VIII

Subsequent Changes in Firm Fundamentals Conditional on Current Bond Rating Changes

Table VIII presents subsequent bond rating actions and subsequent changes in return on equity (ROE), conditional on a previously observed bond rating action for parent firms for 1970 to 1997. For this table, a subsequent bond rating change is defined as a bond rating change within 12 months of the original bond rating change. ROE is measured as income before extraordinary items for the particular year (COMPUSTAT item 18) scaled by the beginning of the year book value of stockholder's equity (COMPUSTAT item 60), while ΔROE is current year's ROE less the prior year's ROE. Firms with negative book values are eliminated from the ROE calculations.

Panel A: Subsequent Bond Ratings Actions Conditional on Original Bond Rating Action

	Subsequent Bond Rating Action			
	No Change	Upgrade	Downgrade	Total
Original bond rating action				
Upgrade %	89.5%	8.3%	2.2%	100%
Number of obs.	772	72	19	863
Downgrade %	73.5%	1.7%	24.8%	100%
Number of obs.	1,129	26	381	1,536

Panel B: Subsequent Changes in Profitability Conditional on Original Bond Rating Action

	Mean (t-statistic)	Standard Deviation	5%	25%	50% (p-value)	75%	95%
Upgrade firms	0.022 (1.670)	0.357	−0.139	−0.003	0.003 (0.013)	0.033	0.214
Downgrade firms	−0.895 (−2.521)	11.96	−1.438	−0.175	−0.030 (0.000)	0.022	0.338

A. Do Stock Prices Reflect Correctly the Implications of Current Bond Ratings Changes for Future Firm Fundamentals?

Table VIII illustrates the relation between current ratings changes and future ratings and earnings changes for the parent firms sample. Future ratings changes are defined as those that occur within one year of the current ratings changes. Earnings changes are defined as changes in return-on-equity (ROE) in the year subsequent to the rating change.[17] An examination of Panel A of Table VIII reveals a strong relation between current ratings changes and future ratings changes. If a subsequent bond rating action is independent of the prior rating action, then the proportion of upgrades and

[17] ROE is measured as income before extraordinary items (COMPUSTAT item 18) scaled by the beginning of the year book value of stockholder's equity (COMPUSTAT item 60).

downgrades following either rating action should be close to 2:3, the average ratio in our sample. However, rating changes are strongly positively correlated, especially for downgrades. The proportion of upgrades to downgrades following an original upgrade is 4:1. Following a downgrade, the proportion of upgrades to downgrades is a remarkable 1:15. Another noteworthy statistic in Panel A is that nearly 25 percent of downgraded firms receive a second downgrade within the following 12 months. Because downgrades are generally preceded by substantial negative abnormal returns, the fact that one in four downgrades experiences another downgrade within one year helps to illustrate the intuition that downgrades signal bad news about the future.

Panel B in Table VIII also reveals that bond rating changes are a strong signal about changes in future profitability. As one might expect, upgrade (downgrade) firms experience an improvement (a deterioration) in profitability over the one-year period following the rating changes. However, consistent with the pattern observed for stock returns, the change in future earnings is considerably more pronounced for downgrade firms. The mean (median) change in future ROE is 2.2 (0.3) percent for upgrade firms, as compared to −89.5 (−3) percent for downgrade firms.[18] Additional tests reveal that, as one might expect, the relation between current bond ratings changes and changes in future fundamentals is even stronger for small, non-investment-grade firms (results not presented).

The evidence in Table VIII suggests that bond rating changes predict future rating and earnings changes. However, these changes in future fundamentals can only account for the abnormal returns following ratings changes if they are not fully anticipated by the market. We examine the pattern of stock returns at the four subsequent earnings announcements to investigate whether the predictable changes in future fundamentals are properly anticipated. Our motivation is that almost all aspects of firm performance are eventually reflected in earnings, and earnings is probably the most closely followed investment variable. Negative stock returns at subsequent earnings announcements would indicate that the market is surprised by the predictable deteriorations of earnings following downgrades, and would support the market imperfections explanation. The major advantage of this approach is that the short-window return specification minimizes the chance that unusual returns at the earnings announcements might be due to inadequate control for systematic risk rather than to the news in the announcements (for similar arguments and tests, see for example Bernard and Thomas (1990), and LaPorta et al. (1997)).

Table IX, Panel A presents mean announcement-period returns for the sample of parent upgrades and downgrades. We present quarterly announcement returns for the first four quarterly earnings announcements after the

[18] The −89.5 percent mean is clearly influenced by extreme negative realizations. However, an examination of the rest of the empirical distribution of ROE changes confirms that on average downgrades are followed by large deteriorations in profitability.

rating change, both separately for each quarter and cumulated for the year. Quarterly earnings announcement returns are measured as buy-and-hold raw returns over the five-day announcement window minus the value-weighted market index return over the same time interval.[19] Panel A reveals two major results. First, announcement returns have no reliable direction for upgrade firms. Second, for downgrade firms the announcement returns are significantly negative in the first quarter, and at the all-quarters level. Thus, the parent-firm sample offers a somewhat limited support for the conjecture that investors are surprised by the predictable deterioration of earnings following downgrades.

The fairly weak results of Panel A of Table IX could also be a reflection of low test power. In search for more power, we present the earnings announcement returns for small non-investment-grade parent issues in Panel B of Table IX. Recall that both abnormal returns and the relation to future fundamentals are more pronounced for small non-investment-grade debt firms. The evidence in Panel B again reveals no abnormal announcement returns following upgrades. However, consistent with expectations, the pattern of negative earnings announcement returns is more pronounced for the small non-investment-grade firms. The absolute magnitude of the earnings announcement returns for all quarters is about −3 percent in Panel B versus about −1.5 percent for Panel A. In addition, the earnings announcement returns in Panel B are significantly negative in both the first and the second quarter following downgrades.

One explanation for the pattern of announcement returns observed in Panels A and B of Table IX is the pattern of total abnormal returns documented earlier. Recall that the underperformance following downgrades is most pronounced in the first months and quarters after the announcement. Thus, it is not entirely surprising that the earnings announcement returns are only significantly negative in the first one or two quarters. However, an examination of Panel B offers another piece of evidence that is important in interpreting the earnings announcement results. For upgrade firms, the number of available observations drops off only slightly from the first quarter to the fourth quarter, from 221 to 210 observations. However, for downgrade firms there is a dramatic drop-off in the number of available earnings announcements, from 669 to 561 observations. The asymmetrically large drop-off for downgrade firms is likely due to the fact that small, non-investment-grade downgrades experience a variety of difficulties, culminating in possible bank-

[19] The quarterly earnings dates are primarily obtained from COMPUSTAT's Quarterly Industrial tape. However, as discussed earlier, COMPUSTAT data are often missing for smaller and otherwise marginal firms where the abnormal returns are likely to be the most pronounced. Therefore, we complement the COMPUSTAT data with hand-collected data on earnings announcements from the Wall Street Journal and the Dow Jones News Retrieval Service. Because the Wall Street Journal and the Dow Jones News Retrieval Service earnings announcement date is often a day later than the actual earnings announcement, we use a five-day window defined as (−3,+1) around each quarterly earnings announcement date (day 0) to assure that we capture the announcement effect.

Table IX

The Magnitude of Earnings Announcement Returns Following Bond Upgrades and Downgrades

This table presents mean market-adjusted stock returns of upgraded and downgraded firms over the subsequent four quarterly earnings announcement periods for parent firms during the period of 1970 to 1997. Quarterly earnings announcement dates are from the COMPUSTAT Quarterly Industrial tape, the *Wall Street Journal* and the *Dow Jones News Retrieval Service*. Announcement returns are measured as the buy-and-hold returns earned over the five-day window (−3, +1) surrounding each earnings announcement (date 0) less the value-weighted market index return over the same time interval. Mean returns for a particular quarter represents the average announcement return for those firms with returns available for that quarter. The total earnings announcement return for each firm (i.e., All Quarters) equals the sum of the individual quarterly earnings announcement returns. If announcement returns are not available for all four quarters, the total announcement return equals the sum of announcement returns over the available dates. The mean "All Quarters" return for each portfolio is the average of these firm-specific total earnings announcement returns. t-statistics are in parentheses. N is the number of available observation at each quarterly announcement. Panel A presents the announcement returns for the parent-issues sample. Panel B presents the announcement returns for the small, non-investment-grade parent issues. Panel C presents additional data on the characteristics of small non-investment-grade downgrades which miss earnings announcements after the first quarter. Performance delistings in Panel C are defined as in Shumway (1997), comprising CRSP delisting codes 500, 520, 551–574, 580, 584.

	First Quarter	Second Quarter	Third Quarter	Fourth Quarter	All Quarters
	Mean Five-Day (−3,+1) Returns				
	Panel A: Earnings Announcement Returns for Parent Firms				
Upgrade firms	0.0028	0.0031	0.0030	−0.0036	0.0051
(t-statistic)	(1.508)	(1.590)	(1.490)	(−1.935)	(1.274)
N	843	822	822	822	843
Downgrade firms	−0.0101	−0.0033	−0.0034	0.0021	−0.0145
(t-statistic)	(−2.774)	(−1.155)	(−1.206)	(0.832)	(−2.779)
N	1449	1379	1342	1321	1449
	Panel B: Earnings Announcement Returns for Small, Non-investment-grade Parent Issues				
Upgrade firms	0.0039	0.0006	0.0021	−0.0033	0.0033
(t-statistic)	(1.102)	(0.093)	(1.097)	(0.090)	(0.312)
N	221	211	210	210	221
Downgrade firms	−0.0183	−0.0128	−0.0058	0.0042	−0.0315
(t-statistic)	(−2.968)	(−2.242)	(−0.482)	(1.431)	(−3.178)
N	669	608	587	561	669

(*continued*)

ruptcies and exchange delistings. If that is true, then the data in Panel B would suffer from a survivorship bias that understates both the total negative abnormal returns and the negative earnings announcement returns following downgrades.

Table IX—*Continued*

Panel C: The Characteristics of Small, Non-investment-grade Parent Issues Missing At Least One Subsequent Earnings Announcement

C1: Announcement Returns

	First Quarter	Second Quarter	Third Quarter	Fourth Quarter	All Quarters	One-year CAR
Returns	−0.0257	−0.0533	−0.0477	—	−0.0751	−0.4183
N	108	66	38	0	108	108

C2: Bankruptcies and Exchange Delistings

Bankrupt
49

Not bankrupt
59

Total number of observations missing at least one subsequent earnings announcement
108

Performance Delisting
65

No performance delisting
43

Available Delisting Return
13

No Delisting Return
52

Panel C in Table IX presents more specific evidence on the survivorship bias for the 108 small non-investment-grade downgrades in Panel B that are missing at least one earnings announcement.[20] The evidence in Panel C confirms the intuition that most downgrades with missing earnings announcements likely experience serious financial difficulties, falling stock prices, and even business failure. The available earnings announcement returns for this troubled group are considerably more negative than those for the aver-

[20] There are 108 firms that are missing at least one earnings announcement return for small, non-investment-grade parent issues, and 128 such firms for the sample of all parent issue. Thus, the small non-investment-grade subsample accounts for the majority of parent observations that miss at least one earnings announcement return. Therefore, the results for the observations with missing earnings announcement returns are quite similar across the full parent sample and the small non-investment-grade subsample.

age downgrade in Panel B of Table IX. In addition, the one-year CAR for this group is about -0.42, substantially more negative than the corresponding CAR of -0.21 for the average small non-investment-grade issue (from Panel B in Table VII). This evidence confirms the conjecture that the results in Panels A and B in Table IX suffer from a survivorship bias that understates the magnitude of the negative announcement returns, especially for later quarters.

The evidence in Panel C also reveals that most of the firms missing earnings announcements enter bankruptcy and/or experience exchange delistings for performance reasons. Using the *Wall Street Journal* and the *Dow Jones News Retrieval Service*, we could verify that 49 of the total 108 firms entered bankruptcy or liquidation procedures. In addition, the CRSP delisting tape identifies 65 of the 108 firms as delisted for performance reasons, where 52 of these 65 performance delistings have no delisting returns.[21] This last fact is important because it is consistent with Shumway (1997), who shows that most performance delistings in CRSP miss delisting returns. Shumway (1997) also finds that most omitted delisting returns are substantially negative. Thus, because downgrades are often followed by performance delistings, the omission of the delisting returns in the CRSP data results in an upward bias in the calculated abnormal returns following downgrades. In other words, the performance delisting bias in CRSP likely leads to an understatement of the negative abnormal returns following downgrades.[22]

Summarizing, returns at subsequent earnings announcements indicate that investors are surprised by the predictable deteriorations in earnings following downgrades. The negative earnings announcement returns are only detectable in the first and second subsequent quarter, but that might be at least partly due to a survivorship bias in computing the announcement returns for later quarters.

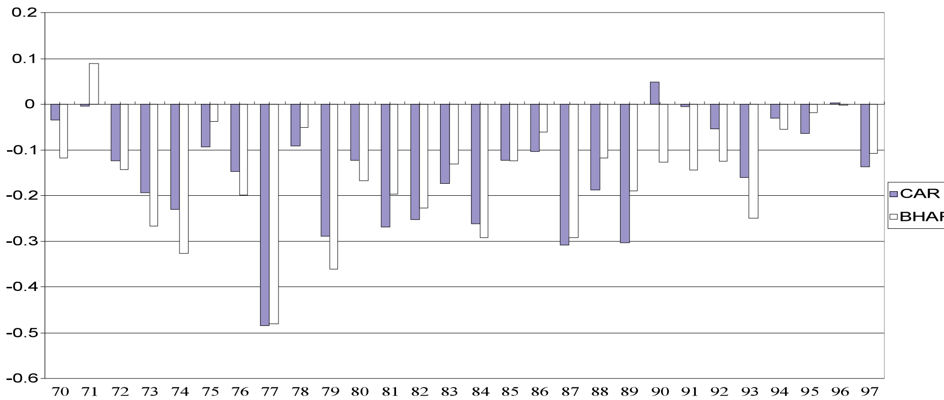
B. Are the Poor Returns Following Downgrades Persistent?

As mentioned before, we also examine whether the average underperformance of downgrade firms is also persistent. If downgrade firms earn low returns because they provide a hedge against some type of systematic risk, then this hedge should result in positive returns in at least some periods. Finding persistently negative abnormal returns would cast doubt on a systematic-risk explanation for the underperformance.

[21] Delisting for performance reasons is defined as in Shumway (1997). Essentially, performance delistings comprise bankruptcy and liquidations delistings along with delistings for a host of other poor performance reasons such as insufficient capital, insufficient number of market-makers, failure to pay fees, and so forth.

[22] This understatement is likely material. Shumway (1997) estimates that the missing performance delisting returns are on the magnitude of at least -30 percent. Multiplying the 53 observations with missing performance delisting returns by the -30 percent estimate and dividing by the 669 downgrades in Panel B of Table IX yields about -2.5 percent. This result implies that negative portfolio CARs for parent small non-investment-grade downgrades are understated by about 2.5 percent.

Panel A: The Distribution of CARs and BHARs Following Parent Downgrades Over Time



Panel B: The distribution of CARs and BHARs following parent downgrades by aggregated time periods

	1970-1978	1979-1987	1988-1997
CARs	-0.2065	-0.2044	-0.0809
t-stat.	(-3.490)	(-7.899)	(-4.066)
N	85	505	946
BHARs	-0.2149	-0.1889	-0.1110
t-stat.	(-4.316)	(-8.230)	(-5.711)
N	85	505	946

Figure 2. The distribution of abnormal returns following parent downgrades over time. Panel A presents the yearly distribution of CARs and BHARs following downgrades for parent firms from 1970 to 1997. Panel B presents the abnormal returns following downgrades aggregated in three time periods, 1970 to 1978, 1979 to 1987, and 1988 to 1997.

Panel A in Figure 2 presents the distribution of one-year CARs and BHARs for parent-firm downgrades by calendar year. CARs are negative in 26 out of 28 years. In addition, the two positive CARs are small, one of about 3 percent and one of 5 percent. BHARs are negative in 27 out of 28 years. The only positive BHAR is fairly small, about 9 percent, and it occurs in 1971, a year with a total of nine observations. Panel B in Figure 2 presents CAR and BHAR results aggregated by three approximately equal-length periods, 1970 to 1978, 1979 to 1987, and 1988 to 1997. An examination of Panel B reveals large and significantly negative CARs and BHARs for all periods. In addition, Panel B shows no reliable pattern in the underperformance over time, although the negative abnormal returns seem less extreme in the 1990s.

Summarizing, the evidence in Figure 2 suggests that the poor returns following downgrades are strongly persistent, which casts further doubt on a systematic-risk explanation. The strong persistence of abnormal returns over nearly 30 years also questions the period-specific explanation, although the tapering off of abnormal returns in the 1990s could indicate that investors are learning from the past poor returns following downgrades.

V. Conclusion

This study examines the long-run stock returns following bond ratings changes. We use a comprehensive sample that comprises essentially all Moody's bond rating changes during 1970 to 1997. We examine both cumulative abnormal returns and buy-and-hold returns, and control for size and book-to-market. We find no reliable abnormal returns following upgrades, whereas there are substantial negative abnormal returns following downgrades. The underperformance is most pronounced in the first months following downgrades, lasts at least a year, and is on the magnitude of -10 to -14 percent at the one-year horizon. The poor returns of downgrade firms are more pronounced for small and low-credit-quality firms. In search for explanations, we document that downgrades are strong predictors of future deteriorations in earnings. More importantly, we find evidence of negative stock reactions at the announcements of these predictable future deteriorations in earnings. In addition, downgrade firms earn poor returns not only on average but in nearly all years of our sample, which further questions a systematic-risk explanation.

The evidence in this paper also opens new questions. An interesting pattern that emerges from this study and other research is that abnormal returns tend to be more extreme "on the downside." For example, Bernard and Thomas (1989) find that the post-earnings-announcement drift is more pronounced for negative earnings surprises. Michaely, Thaler, and Womack (1995) find more extreme abnormal returns following dividend omissions and cuts, as opposed to dividend increases and initiations. A potential research question that emerges from these findings is whether the abnormal returns "on the downside" are more extreme because there are different information implications of good versus bad news, and investors fail to recognize this fact. Another possible explanation is that well-known information-processing biases like optimism might also play a role. Perhaps optimism naturally results in more erroneous conclusions when applied to fairly negative situations, and thus negative situations are followed by more extreme subsequent adjustments in stock prices.

REFERENCES

Altman, Edward I., 1993, *Corporate Financial Distress and Bankruptcy*, 2nd ed. (Wiley, New York).

Barber, Brad M., and John D. Lyon, 1997, Detecting long-run abnormal stock returns: The empirical power and specification of test statistics, *Journal of Financial Economics* 43, 341–372.

Bernard, Victor L., and Jacob K. Thomas, 1989, Post-earnings announcement drift: Delayed price response or risk premium, *Journal of Accounting Research* 27, 1–36.

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

Brav, Alon, and Paul A. Gompers, 1997, Myth or reality? The long-run underperformance of initial public offerings: Evidence from venture and nonventure capital-backed companies, *Journal of Finance* 52, 1791–1821.

Blume, Marshall E., Felix Lim, and A. Craig MacKinlay, 1998, The declining credit quality of U.S. corporate debt: Myth or reality?, *Journal of Finance* 53, 1389–1413.

Chan, Louis, Narasimhan Jegadeesh, and Josef Lakonishok, 1996, Momentum strategies, *Journal of Finance* 51, 1681–1713.

Dichev, Ilia D., 1998, Is the risk of bankruptcy a systematic risk?, *Journal of Finance* 53, 1131–1148.

Fama, Eugene F., 1998, Market efficiency, long-term returns, and behavioral finance, *Journal of Financial Economics* 49, 283–306.

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

Foster, George, Chris Olsen, and Terry Shevlin, 1984, Earnings releases, anomalies, and the behavior of security returns, *Accounting Review* 59, 574–603.

Glascock, John L., Wallace N. Davidson, and Glenn V. Henderson, 1987, Announcement effects of Moody's bond ratings changes on equity returns, *Quarterly Journal of Business and Economics* 26, 67–78.

Hand, John R. M., Robert W. Holthausen, and Richard W. Leftwich, 1992, The effect of bond rating agency announcements on bond and stock prices, *Journal of Finance* 47, 733–752.

Holthausen, Robert W., and Richard W. Leftwich, 1986, The effect of bond rating changes on common stock prices, *Journal of Financial Economics* 17, 57–89.

Ikenberry, David, Josef Lakonishok, and Theo Vermaelen, 1995, Market underreaction to open market share repurchases, *Journal of Financial Economics* 39, 181–208.

Ikenberry, David, Graeme Rankine, and Earl K. Stice, 1996, What do stock splits really signal?, *Journal of Financial and Quantitative Analysis* 31, 357–377.

Jaffe, Jeffrey, 1974, Special information and insider trading, *Journal of Business* 47, 411–428.

Jewell, Jeff, and Miles Livingston, 1998, Split ratings, bond yields, and underwriter spreads, *Journal of Financial Research* 21, 185–204.

Kaplan, Robert, and G. Urwitz, 1979, Statistical models of bond ratings: A methodological inquiry, *Journal of Business* 52, 231–261.

Keenan, Sean C., Lea V. Carty, and Igor Shtogrin, 1998, *Historical Default Rates of Corporate Bond Issuers (1920–1997)*, (Moody's Global Credit Research, New York.)

Kothari, S. P., and Jerold B. Warner, 1997, Measuring long-horizon security price performance, *Journal of Financial Economics* 43, 301–339.

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

LaPorta, Raphael, Joseph Lakonishok, Andrei Shleifer and Robert Vishny, 1997, Good news for value stocks: Further evidence on market efficiency, *Journal of Finance* 52, 859–874.

Loughran, Tim, and Jay R. Ritter, 1995, The new issues puzzle, *Journal of Finance* 50, 23–51.

Mandelker, Gershon, 1974, Risk and return: The case of merging firms, *Journal of Financial Economics* 1, 303–335.

Michaely, Roni, Richard Thaler, and Kent Womack, 1995, Price reactions to dividend initiations and omissions: Overreaction or drift?, *Journal of Finance* 50, 573–608.

Pinches, George E., and J. Clay Singleton, 1978, The adjustment of stock prices to bond rating changes, *Journal of Finance* 33, 29–44.

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

Spiess, D. Katherine, and John Affleck-Graves, 1995, Underperformance in long-run stock returns following seasoned equity offerings, *Journal of Financial Economics* 38, 243–267.

Spiess, D. Katherine, and John Affleck-Graves, 1999, The long-run performance of stock returns following debt offerings, *Journal of Financial Economics* 54, 45–73.

Weinstein, Mark I., 1977, The effect of a rating change announcement on bond price, *Journal of Financial Economics* 5, 329–350.