An Investigation of Certain Accounting-Related Stock Market Anomalies

by

Soh Yung Kim

M.S., State University of New York at Buffalo, 2004
B.B.A., Seoul National University, 2001

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate Studies
(Business Administration)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)

April 2012

Abstract

In financial markets, anomalies refer to empirical regularities in which security returns deviate from what would be expected in an informationally efficient market. This dissertation investigates explanations for stock market anomalies related to accounting information as documented by Dichev (1998) and Piotroski (2000).

Using Ohlson's (1980) measure of bankruptcy risk (O-Score), Dichev (1998) documents a bankruptcy risk anomaly in which firms with high bankruptcy risk earn lower than average returns. My study first demonstrates that the negative association between bankruptcy risk and returns does not generalize to alternative measures of bankruptcy risk. Then, by examining the nine individual components of O-Score, I find that funds from operations (FFO) is the only component that is associated with returns. Furthermore, I show that the return-predictive power of FFO is due to cash flows from operations. Taken as a whole, this study provides evidence that Dichev's bankruptcy risk anomaly is a manifestation of investors' under (over)-pricing of cash flows (accrual) component of earnings, i.e., the accrual anomaly documented by Sloan (1996).

The second study investigates the effects of two potentially problematic research design choices which are often made in accounting-based studies of anomalies. I explore these issues by re-examining the results in Piotroski (2000), who finds that a simple, financial statement-based heuristic, when applied to a subset of firms with high book-to-market ratios, can discriminate between the firms that will eventually provide high returns and those that will be poor perform-

ers. I find that the relationship between Piotroski's fundamental signals and subsequent returns is partly driven by the choice of return accumulation periods and the use of equally weighted returns. When the research design controls for both problems, the relationship disappears. Because the methods used in Piotroski are typical of those often employed in the accounting literature, this study suggests that evidence of profitable trading strategies and market inefficiency in the literature is likely to be overstated.

Preface

A version of Chapter 2 in this dissertation has been accepted for publication at *Review of Accounting Studies* [Kim, S. 2012. What is behind the magic of O-Score? An alternative interpretation of Dichev's (1998) bankruptcy risk anomaly, *Review of Accounting Studies*, Forthcoming].

Table of contents

List of tables

List of figures

Acknowledgements

I offer my enduring gratitude to the faculty, staff, and my fellow students at the UBC, who have inspired me to continue my journey of learning. I owe particular thanks to Professors Sandra Chamberlain and Kin Lo, whose penetrating questions taught me to question more deeply. I thank Professor Gerald (Jerry) Feltham for providing theoretical training and for being an excellent example as a researcher, a teacher, and a father. I also thank Professors Joy Begley, Kai Li, and Dan Simunic for their advice on my research and support toward completion of PhD degree. Special thanks are owed to my parents and parents-in-law, who have supported me throughout my lengthy years of education, both morally and financially.

To my wife Sue,

for being a patient endurer, for being a candle flame in dark times, and for being there with me.

To my children Christine, Elizabeth, and Kenneth,

for being proud of me for who I am.

I love you!

Chapter 1: Introduction

In financial markets, anomalies refer to empirical regularities in which security returns deviate from what would be expected in an informationally efficient market. This dissertation investigates explanations for stock market anomalies related to accounting information as documented in two separate studies by Dichev (1998) and Piotroski (2000). This chapter briefly reviews the existing literature and highlights where my studies fit in with the existing literature.

Fama (1970) defines an efficient market as one in which "security prices fully reflect all available information (p. 383)." Whether securities markets are informationally efficient is of great interest to investors, standard setters, researchers, and other market participants because securities prices determine the allocation of wealth among individuals and the allocation of scarce resources among firms (Kothari, 2001). Not surprisingly, there is a large amount of literature in finance and accounting that investigates whether information appears to be efficiently impounded into prices.

Early capital market studies that examined security prices around economic events such as earnings announcements (Ball and Brown, 1968; Beaver, 1968) or stock splits (Fama, Fisher, Jensen, and Roll, 1969) largely support market efficiency. However, later studies with more refined data challenge market efficiency. For example, studies on stock price reactions to earnings announcements find that returns drift follows the initial price reaction. Foster, Olsen, and Shevlin (1984) and Bernard and Thomas (1990) document that abnormal returns following earnings announcement last up to a year and that the magnitude of the effect is both statistically and economically significant for the extreme good and bad earnings news portfolios, suggesting that the market underreacts to information in earnings announcements. More recently, Sloan (1996) challenges market efficiency with what has been termed as the ‘accrual anomaly.’ His evidence suggests that the market assigns a higher persistence to the accrual component of earnings and over (under)-prices firms with earnings containing high (low) accrual components. As a result, an investment strategy that exploits this mispricing earns abnormal returns of around 10 percent in the year following the release of financial statements. Sloan suggests that the market fails to correctly incorporate the information in accruals into stock prices.

Although long-horizon tests of market efficiency report economically large abnormal returns, not all researchers accept these returns as evidence of market inefficiency. For example, Kothari (2001) emphasizes that long-horizon tests suffer from several data problems such as risk adjustments, sample selection biases, and inference problems arising from cross-sectional dependence. A number of studies (e.g., Fama, 1998; Guay, 2000) argue that studies purporting to identify market inefficiencies typically do not provide a plausible and rejectable alternative hypothesis to market efficiency. With the exception of a few studies (e.g., Bernard and Thomas, 1990; Sloan, 1996), most studies that show evidence of inefficiency fail to provide a theory that specifies conditions under which market under- and over-reaction is forecasted.[1] Guay remarks, "Rather than providing evidence in support of a pricing theory, they are marketed as descriptive evidence of a 'puzzle', to be resolved by future research (p. 50)." This statement precisely describes the current status of the bankruptcy risk anomaly literature and for that reason, Chapter 2 of this dissertation investigates this anomaly.

The bankruptcy risk anomaly, documented by Dichev (1998), refers to the empirical regularity in which a composite measure of bankruptcy risk predicts future returns. Using Ohlson's (1980) measure of bankruptcy risk (O-Score), Dichev finds that firms with high bankruptcy risk earn lower than average returns during his sample period of 1981 to 1995. As noted in Dichev, his original expectation was to find the opposite relationship based on the conjecture that higher risk, more insolvent firms should have higher average subsequent returns compared with less insolvent firms. As a result of his findings, Dichev suggests that market prices do not fully impound the implications of available bankruptcy information.

Dichev (1998) is a seminal work in this literature; it is the first study that investigates the relationship between bankruptcy risk and returns. Prior to Dichev, the focus of bankruptcy studies had been on the development of models to better predict future bankruptcy (e.g., Beaver,

---

[1] There is a growing body of research that comes under the heading of "limited attention" that appears to offer such a theory (e.g., Hirshleifer and Teoh, 2003; Hirshleifer et al., 2008; Hirshleifer, Lim, and Teoh, 2009; Hirshleifer, Lim, and Teoh, 2011).

1966; Altman, 1968; Ohlson, 1980; Zmijewski, 1984). Dichev's counterintuitive results led to numerous follow-up studies. However, use of allegedly better measures of bankruptcy risk and different samples in these studies make comparisons problematic.

An example of these comparability issues can be seen in Vassalou and Xing (2004) and Campbell, Hilscher, and Szilagyi (2008), two often cited follow-up studies. Interestingly, both studies were published in *The Journal of Finance*, but they reach different conclusions. The former, using a market-based measure of bankruptcy risk, Default Likelihood Indicator (DLI), derived from Merton's (1974) framework, finds evidence that firms with high bankruptcy risk earn higher returns. By contrast, the latter, using its own measure of bankruptcy risk based on accounting information as well as market information, finds that stocks with high bankruptcy risk have delivered low returns since 1981. Subsequent studies show either a positive or a negative relationship between their measures of bankruptcy risk and returns (e.g., Chava and Purnanadam, 2010; George and Hwang, 2010). Thus, the literature seems far from settled with respect to the relationship between bankruptcy risk and returns.

To help understand these mixed results and the bankruptcy risk anomaly, I analyze a single sample with two different proxies for bankruptcy risk: Ohlson's (1980) O-Score, the indicator of risk used by Dichev (1998), and BSM-Prob, a DLI measure used in Hillegeist et al. (2004). Using O-Score, I successfully replicate the results of Dichev but, using BSM-Prob, I am unable to replicate his results. Next, I investigate the source of O-Score's predictive ability by examining the individual components of O-Score. I find that funds from operations (FFO) is the only component that is associated with returns. Furthermore, I show that the return-predictive power of FFO is due to cash flows from operations. Taken as a whole, Chapter 2 provides evidence that Dichev's bankruptcy risk anomaly is a manifestation of investors' under (over)-pricing of cash flows (accrual) component of earnings, i.e., Sloan's (1996) accrual anomaly.

The bankruptcy risk anomaly and the accrual anomaly are only two examples of anomalies documented in the literature. As noted in Keim (2008), the number of documented anomalies is large and growing. While anomaly studies report economically significant anomalous returns, many other studies (e.g., Malkiel, 1995; Carhart, 1997) document that, on average, managers of mutual funds and other actively managed funds do not outperform the market. If a primary objective of fund managers is to invest their capital in assets that generate the highest possible risk-adjusted returns, then why are returns earned by fund managers not as great as returns documented by academic researchers? Moreover, how does documented mispricing persist without being arbitraged away by fund managers?[2] These questions motivate Chapter 3 of this dissertation, where I investigate the implementability of trading strategies based on accounting information.

The chapter investigates the effects of two potentially problematic research design choic-

[2] Some work has been done to address this question. For example, Mashruwala, Rajgopal, and Shevlin (2006) explain Sloan's (1996) accrual anomaly using idiosyncratic risk and transaction costs.

es that are often made in accounting-based studies of anomalies. Both of these research design choices are relevant to the class of studies that employ zero-investment hedge portfolios. The two issues are: 1) the use of equal- versus value-weights in forming the portfolios and 2) the selection of return accumulation periods based on firm specific fiscal year-ends, rather than establishing a common investment date for all firms in the two hedge portfolios.

Fama and French (2008) argue that equal weighting of stocks within the hedge portfolios can be called into question because of real-world implementation issues. For example, researchers that use equal-weighted portfolios assume that sufficient quantities of small-cap stocks can be acquired or shorted, contrary to real-world conditions. Despite this, many accounting studies of anomalies use equal-weighted portfolios and do not provide evidence that their results are robust to an alternative portfolio weighting scheme. Secondly, accounting researchers sometimes employ a return-aggregation approach that cannot be implemented by money managers. If a money manager were to sort firms on the variable of interest such as book-to-market, she cannot determine the total number of firms in each side of the hedge portfolio until the accounting information of the entire portfolio becomes available. However, in forming portfolios shortly after the fiscal year-end, accounting researchers assume knowledge about the identity and weighting of firms that could not be known to money managers. I refer to this issue as the problem of unknown portfolio weights.

I explore these two methodological issues by re-examining the results in Piotroski (2000).

Piotroski finds that a simple, financial statement-based heuristic, when applied to a subset of firms with high book-to-market ratios, can discriminate between the firms that will eventually provide high returns and those that will perform poorly. In documenting this evidence, Piotroski's research design is subject to both of the aforementioned methodological issues. His study is worth revisiting for two reasons. First, his research question leads him to employ a subset of firms that are likely to be particularly vulnerable to the implementation issues that surround equally weighting firms in hedge portfolios. Second, the research in finance that originally documents excess returns to investing in firms with high book-to-market ratios aggregates returns using a common calendar date (e.g., July 1st) rather than relying on fiscal year-ends to define the return aggregation period as in Piotroski. Thus, the question arises, would Piotroski have obtained similar results even if he had followed the approach used in the finance studies upon which he relied?

I find that the relationship between Piotroski's (2000) fundamental signals and subsequent returns is partly driven by the choice of return accumulation periods and the use of equally weighted returns. When the research design controls for both problems, the relationship disappears. Because the methods used in Piotroski are typical of those often employed in the accounting literature, this study suggests that evidence of profitable trading strategies and market inefficiency in the literature is likely to be overstated.

The remainder of the dissertation is structured as follows. In Chapter 2, I re-examine the bankruptcy risk anomaly and provide an alternative interpretation of Dichev's (1998) findings. In Chapter 3, I investigate the impact of the difference in return accumulation periods and different weighting schemes on the evidence of market efficiency by replicating and extending the analysis in Piotroski (2000). In the final chapter, I summarize the findings from prior chapters and outline major contributions made by this dissertation.

Chapter 2: What is behind the magic of O-Score? An alternative interpretation of Dichev's (1998) bankruptcy risk anomaly[3]

2.1 Introduction

The bankruptcy risk anomaly documented by Dichev (1998) refers to the empirical regularity in which a composite measure of bankruptcy risk predicts future returns. Using Ohlson's (1980) measure of bankruptcy risk (O-Score), Dichev finds that firms with high bankruptcy risk earn lower than average returns during his sample period of 1981−1995. As noted in Dichev, his original expectation was to find the opposite relationship based on the conjecture that "if the risk of bankruptcy is at least partly systematic, then more insolvent firms should have higher average subsequent returns as compared to less insolvent firms (p. 1134)." As a result of his findings, Dichev suggests that market prices do not fully impound the implications of available bankruptcy information.

Subsequent studies using different measures of bankruptcy risk and samples provide mixed evidence. A group of researchers that use Ohlson's (1980) O-Score as a proxy of bankruptcy risk supports Dichev's (1998) findings. For example, Griffin and Lemmon (2002) find that the negative relationship between O-Score and returns is strong among firms with low book-to-market ratios (i.e., growth firms), suggesting that the bankruptcy risk anomaly is likely to

[3] A version of this chapter has been accepted for publication at *Review of Accounting Studies*.

occur as a result of a high degree of information asymmetry among growth firms, and George and Hwang (2010) assert that the negative relationship between returns and leverage is responsible for the negative association between bankruptcy risk and returns. By contrast, another group of researchers, relying on the other measures of bankruptcy risk, challenges Dichev's conclusion by showing a positive relationship between bankruptcy risk and returns (e.g., Vassalou and Xing, 2004; Chava and Purnanandam, 2010) or no reliable relationship between the two (e.g., Anginer and Yildizhan, 2010).[4] Although the latter group of studies refute Dichev's conclusion, the causes of Dichev's findings have not been answered. This study attempts to fill this gap.

To help understand these mixed results and the bankruptcy risk anomaly, I analyze a single sample with two different proxies of bankruptcy risk: O-Score and BSM-Prob (Hillegeist et al., 2004). I choose to use BSM-Prob because Hillegeist et al. empirically show that their measure is superior to O-Score in terms of bankruptcy prediction. Using two distinct proxies increases the ability to identify the source of the anomaly and employing two proxies within the identical research setting ensures that my results are not affected by differences in sample periods or composition. With a dataset spanning 1978 to 2007, I find a reliable negative association between O-Score and returns, but fail to find any association between BSM-Prob and returns. These results

[4] One exception is Campbell, Hilscher, and Szilagyi (2008), where they find that stocks with high bankruptcy risk have delivered low returns since 1981. Their (own) measure of bankruptcy risk utilizes accounting information as well as market information and lies outside the scope of this study.

suggest that the bankruptcy risk anomaly is construct-dependent: some measures of bankruptcy risk predict future returns while others do not. The question arises: what is the source of O-Score's predictive ability? I tackle this question by examining the individual components of O-Score.

In regressions of returns on the individual components of O-Score, I find that out of the nine accounting variables, funds from operations divided by total liabilities (FFO/TL) is the only component that predicts returns. In a complementary analysis, I show that the zero-investment trading strategy long on firms with the highest decile rank of FFO/TL and short on firms with the lowest decile rank produces economically and statistically positive net returns over the period of 1978–2007. Further investigation reveals that the association between O-Score and subsequent returns disappears once FFO/TL is controlled for. This implies that the return-predictive power of O-Score is driven by the association of returns with FFO/TL, the component that is intended to capture the degree to which the internally generated working capital of a firm is available to satisfy its debt obligations. Given this result, the question becomes: why does FFO/TL predict returns?

Prompted by a close relationship between funds from operations and cash flows from operations, I conjecture that FFO/TL mimics cash flows from operations deflated by average total assets (CFO/ATA). Since Sloan (1996) and Houge and Loughran (2000) find CFO/ATA to be associated with subsequent returns, I conjecture that the ability of FFO/TL to predict returns is not distinct from the ability of CFO/ATA to predict returns. The zero-investment trading strategy long on firms with the highest decile rank of CFO/ATA and short on firms with the lowest decile rank generates returns of a similar or higher magnitude, in comparison to that using FFO/TL. A close comparison of annual net returns from the two trading strategies reveals that (1) signs of net returns from the two strategies agree in most of my sample years, and that (2) when the strategy using CFO/ATA fails to provide positive net returns, the strategy using FFO/TL also fails to produce positive net returns, with few exceptions. Furthermore, the regression analysis demonstrates that once CFO/ATA is controlled for, the association of returns with FFO/TL diminishes, suggesting that the return-predictive power of FFO/TL is related to the association of returns with CFO/ATA. Lastly, this finding is confirmed in the analysis using a two-way sort using quintiles that are independently formed according to the rankings of FFO/TL and CFO/ATA.

Closing the investigation, I predict and find that O-Score is no longer associated with subsequent returns once CFO/ATA is controlled for. The most striking evidence is unveiled in the analysis using quintiles that are independently formed using O-Score and CFO/ATA. Not all firms with high O-Score earn lower than average returns; among firms in the highest quintile of O-Score, only the firms that belong to the lowest quintile of CFO/ATA deliver low returns.

Taken as a whole, my results suggest that the negative association between Ohlson's (1980) measure of bankruptcy risk and subsequent returns is a manifestation of Sloan's (1996) accrual anomaly.[5] Thus, it appears that Dichev's (1998) anomaly is due to investors not fully impounding the implication in the accruals and cash flows components of earnings, rather than their failure to incorporate available bankruptcy information. FFO/TL happens to serve as a bridge between the two seemingly unrelated stock market anomalies.

This study contributes to the existing research in several ways. First, it provides an alternative explanation for Dichev's (1998) findings. In this sense, this study is complementary to the explanations provided by Griffin and Lemmon (2002), which suggest that the bankruptcy anomaly documented by Dichev is likely to occur as a result of a high degrees of information asymmetry among firms with low book-to-market ratios, and George and Hwang (2010), which suggests that the negative relationship between returns and leverage is responsible for the Dichev's results. Second, this study highlights that the choice of bankruptcy risk measure can lead to different conclusions. Third, by relating the bankruptcy risk anomaly to the accrual anomaly, this study contributes to a stream of research that investigates the relationship between anomalies (e.g., Collins and Hribar, 2000; Desai, Rajgopal, and Venkatachalam, 2004).

---

[5] The results in Houge and Loughran (2000) and Pincus, Rajgopal, and Venkatachalam (2007) suggest that there may be an accrual anomaly and a cash flow anomaly and that they do not always occur together, i.e., there is no mechanical relationship going on such that if one is present, the other must also be present. Section 2.4.4 investigates whether the negative association between O-Score and returns would disappear if Sloan's original accrual measure is employed.

2.2 Empirical analysis I: How robust is the bankruptcy risk anomaly when using an alternative measure of bankruptcy risk?

This section investigates the robustness of the bankruptcy risk anomaly using O-Score and BSM-Prob, where the former exclusively uses information from financial statements and the latter extracts the bankruptcy-related information from market prices as well as financial statement data. Employing the two measures in an identical setting ensures that my results are not affected by differences in sample periods or composition. I use BSM-Prob as an alternative proxy because using two distinct proxies increases the ability to identify the source of the anomaly. In addition, Hillegeist et al. (2004) empirically show that their measure is superior to those from Altman's (1968) and Ohlson's (1980) models in terms of bankruptcy prediction. Specifically, Hillegeist et al. employ relative information content tests on out-of-sample predictions and demonstrate that BSM-Prob contains significantly more information about the probability of bankruptcy than any of accounting-based measures: the original Z-Score and O-Score and "updated" Z-Score and O-Score (scores computed using updated coefficient estimates). If there is a bankruptcy risk anomaly, then a better measure of bankruptcy risk should be able to better predict returns. The description of the sample and the definition of the variables are provided in Section 2.2.1, followed by empirical results in Section 2.2.2.

2.2.1 Sample formation and variable measurement

The empirical tests in Section 2.2 are conducted using NYSE/AMEX/NASDAQ firms with all available data in the intersection of the 2008 versions of the Center for Research in Security Prices (CRSP) monthly stock file and the merged COMPUSTAT/CRSP annual fundamental file (the successor to the industrial file). I exclude Canadian firms to avoid having to adjust accounting numbers into a common currency.[6] I further restrict the sample to industrial firms because Ohlson's (1980) models are intended for industrials. Thus, this study only includes firms with CRSP Standard Industrial Classification (SIC) codes 1 to 3999 and 5000 to 5999. I also limit the sample to firms with ordinary shares (CRSP share codes of 10 and 11); i.e., I exclude ADRs, REITs, and units of beneficial interest from my analysis. I restrict the sample to post-1976 years because Ohlson derives his models using accounting data from 1970 to 1976. Furthermore, my research design requires accounting data for year $t-1$ and as a result, the sample period begins in 1978. The sample period ends in 2007 because I require returns data for year $t+1$.

After restricting firm-years to those with sufficient data to compute the variables described below, my final sample is 74,737 firm-years representing 8,616 firms. The analyses in this study use firms with any fiscal year-end month in order to maximize sample size. However, some researchers prefer to restrict their sample to firms with a December 31 year-end to facilitate implementation of the trading strategies (e.g., Lakonishok, Shleifer, and Vishny, 1994; Abarbanell and Bushee, 1998). Thus, I also analyze a sample of firms with a December 31 year-end.

---

[6] Unlike the COMPUSTAT Industrial File, in which all data were displayed in U.S. dollars, the Fundamental File displays data for Canadian companies (dual listed or not) using Canadian dollars.

Griffin and Lemmon (2002) find that the bankruptcy risk anomaly can be traced to firms with a small market capitalization and low book-to-market ratios.[7] To address these alternatives, I examine four subsamples. The small (large) subsamples include those firms with market capitalization smaller (larger) than the first quintile cutoff of NYSE size distribution at the end of June of each year *t*.[8] The growth (value) subsamples include firms whose book-to-market is smaller (larger) than the median of NYSE book-to-market distribution computed at the end of June of each year *t*. I obtain NYSE size break-points and book-to-market break-points from Professor Kenneth French's data library.[9]

Consistent with Dichev (1998), Griffin and Lemmon (2002), and George and Hwang

---

[7] More precisely, Griffin and Lemmon find that the negative association between Ohlson's (1980) measure of bankruptcy risk and subsequent returns is primarily from growth firms, suggesting that the bankruptcy risk anomaly is likely to occur as a result of a high degree of information asymmetry among growth firms. However, I repeat my analysis using small and large firms as well as growth and value firms because growth firms tend to be small and Kothari (2001) documents that evidence of market inefficiency tends to be more pronounced among small firms.

[8] Firms listed in AMEX or NASDAQ tend to be smaller than those in NYSE. Hence, using the median of NYSE size distribution leads me to classify less than 20 percent of firm-years to the subsample of large firms. Thus, I employ the first quintile of NYSE size distribution. In any case, the tenor of results remains unchanged when I use the median as a size cutoff.

[9] The data library is available at http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/.

(2010), the variable *O-Score* is calculated using the original coefficients from Model 1 in Ohlson (1980) as follows:

> *O-Score* =
>
> –1.32 – 0.407*log(total assets/GNP price-level index) + 6.03*(total liabilities/total assets) – 1.43*(working capital/total assets) + 0.076*(current liabilities/current assets) – 1.72*(1 if total liabilities > total assets, else 0) – 2.37*(net income/total assets) – 1.83*(funds from operations/total liabilities) + 0.285*(1 if net loss for the last two years, else 0) – 0.521*(net income_t – net income_{t-1})/ (|net income_t| + |net income_{t-1}|).

The inputs to the calculation of $O\text{-}Score_t$ are from the financial statements of fiscal years ending in calendar year $t-1$, which ensures that *O-Score* is estimated on an *ex ante* basis. All the inputs to the calculation of *O-Score* are readily available from COMPUSTAT, except funds from operations, which are only available up to year 1987. In post-1987 years, I derive funds from operations by subtracting non-cash working capital adjustments from cash flows from operations.

Using Ohlson's (1980) original coefficients implicitly assumes that the association between the accounting variables and the probability of bankruptcy has not changed since Ohlson's original sample period of 1970–1976. However, as noted in Hillegeist et al. (2004, p. 15), the assumption does not necessarily hold. First, the Bankruptcy Reform Act of 1978 made a substantial change in the legal environment, which was followed by an increase in the number of strategic bankruptcy filings. Second, significant intervening changes in accounting rules have also changed the original associations through differences in how the accounting variables are measured. Nevertheless, the (in)validity of this assumption does not pose any problem for the purposes of this study, as my focus is the association of returns with Ohlson's original measure of bankruptcy risk as documented in Dichev (1998), not on how well Ohlson's measure predicts future bankruptcies.[10]

The market-based measure of bankruptcy risk used in this study, *BSM-Prob*, is from Hillegeist et al. (2004), which use the insights from Black and Scholes (1973) and Merton (1974) (BSM). The BSM approach views the firm's equity as a call option on the value of the firm's assets. The measure captures the idea that when the value of a firm's assets is greater than the strike price (i.e., the face value of the liabilities), equity holders find it profitable to continue owning the firm. When the value of the firm's assets is smaller than the strike price at time *T*, the call option is assumed to be unexercised and the equity holders turn over bankrupt firm to the debtholders. The probability of each outcome is an important determinant of the value of the call option, and these probabilities are embedded in the BSM option-pricing model.

Estimation of *BSM-Prob* requires five key inputs, namely common dividends, preferred dividends, total liabilities, market value of equity at the fiscal year-end, and annualized standard deviation of the daily returns. I calculate the annualized standard deviation of a firm's returns by multiplying the daily standard deviation by the square root of 252, the average number of trading

[10] Studies that examine the bankruptcy-predictive power of Ohlson's O-Score include Begley, Ming, and Watts (1996) and Hillegeist et al. (2004).

days in a year. Note that as in the estimation of *O-Score*$_t$, all applicable accounting data for the estimation of *BSM-Prob*$_t$ is derived from the financial statements of fiscal year ending in calendar year $t-1$. Details on the estimation of *BSM-Prob* are outlined in Appendix A.

For each firm-year, I compute subsequent realized returns (*Ret1yr*$_t$), which are defined as the buy-and-hold returns from July of year t to June of year $t+1$. This study requires a minimum six-month gap to ensure that the accounting data is available to calculate bankruptcy measures before it is used to explain the returns. Although Ohlson (1980, Table 2, p. 116) shows that the financial reports of firms that go bankrupt are often delayed, sometimes significantly, I choose a six-month lead time as in Dichev (1998), who supports this choice by noting the evidence in Alford, Jones, and Zmijewski (1994) that 98 percent of the 10-K reports are filed by the end of the fifth month after fiscal year-end. Consistent with prior research, when a firm delists, I use the delisting return in the delisting month. If a firm's delisting is due to liquidation or a forced delisting by either the exchange delisting code or the Securities and Exchange Commission (SEC) (CRSP delisting codes 400 to 699) and the delisting return is missing, the delisting return is set to negative 100 percent.

I also compute measures of asset pricing risk that are shown to be associated with subsequent returns (i.e., firm beta, size, and book-to-market). Controlling for these risk measures is important to the extent that returns and bankruptcy risk are both correlated with these measures. Following Fama and French (1992), *Beta*$_t$ is estimated with the Capital Asset Pricing Model

(CAPM) using monthly returns for at least 24 of 60 months preceding July of year *t*. Size (ME_t) is defined as the log of a firm's market equity in millions at the end of June of year *t*. Book-to-market (BE/ME_t) is defined as book value of common equity for the fiscal year ending in calendar year $t-1$, divided by market value of equity at the end of December of year $t-1$. Instead of *BE/ME*, Fama and French and other studies commonly use log(BE/ME) and exclude firms with negative book value of equity because log(BE/ME) is not defined. However, as in Dichev (1998), I do not remove these firms from the sample because having a negative book value of equity is not a random event; rather, these firms are likely to have a high bankruptcy risk. Note that the positive/negative book value dichotomy is an explicit prediction variable in Ohlson's (1980) model and thus, eliminating firms with negative book value is inappropriate. In any case, the results remain qualitatively unchanged in a sensitivity check eliminating firms with negative book values (1,910 out of 74,737 firm-years) and using log(BE/ME).

2.2.2 Results

To better understand the relationship between Ohlson's (1980) measure of bankruptcy risk and the other variables, I form portfolios according to the rankings of $O\text{-}Score_t$ at the end of June of each year *t* and report the average values of each variable in Panel A of Table 2.1. In Panel A, I also provide the return distribution association with *O-Prob*. *O-Prob* is a monotonic transformation of *O-Score* derived using the logistic cumulative distribution (i.e., $O\text{-}Prob = e^{O\text{-}Score}/(1 + e^{O\text{-}Score}) * 100\%$), for ease of comparison with *BSM-Prob* in Panel B. Patterns of *O-Prob* and *BSM-*

Prob display extreme right skewness, which is expected given the low frequency of actual bankruptcy in the COMPUSTAT population.[11] Note that the choice between *O-Score* and *O-Prob* does not have any impact on the results of the subsequent analyses because I use relative rankings of the variables in each year *t* in the analyses.

In Panel A, firms in the extreme high or low *O-Score* deciles have higher *Beta*. Firms with higher *O-Score* have smaller market capitalization; this is not surprising given that smaller firms are more likely to go bankrupt. Indeed, Shumway (2001) finds that firm size is highly related to bankruptcy. The pattern of *BE/ME* demonstrates that firms with higher *O-Score* tend to have higher book-to-market, with exception of firms in the highest *O-Score* decile, which have the smallest book-to-market, consistent with the pattern shown in Dichev (1998, Table 4, p. 1140).[12] As suggested by Dichev, the reason for this pattern is that, unlike market value which is always positive, the book value of firms with the highest bankruptcy risk can be very small or negative if losses are high enough. The pattern of *Ret1yr* suggests that market prices do not fully

[11] The average annual bankruptcy rate during 1980–2000 period was 0.97 percent among the industrial firms in the intersection of the CRSP and COMPUSTAT databases (Hillegeist et al., 2004, Table 1, p. 12).

[12] A decline of book-to-market from the second highest *O-Score* decile portfolio to the highest decile portfolio is much milder in Dichev because he winsorizes all the variables at the 1st and 99th percentile, whereas I use the variables without any adjustment. I intentionally do not follow Dichev in this respect due to the concern raised by Kothari, Sabino, and Zach (2005), where they demonstrate that data deletion induces a spurious association between subsequent returns and *ex ante* information variables.

impound the implications of available bankruptcy information as proxied by *O-Score*. Although firms with higher bankruptcy risk do not necessarily have lower subsequent returns, firms in the highest *O-Score* decile do have the lowest subsequent returns, followed by the second highest *O-Score* decile having the second lowest subsequent returns. Dichev (Table 6, p. 1145) also reports a similar non-monotonic relationship between *O-Score* and subsequent returns.

With respect to *BSM-Prob*, Panel B shows that firms with higher *BSM-Prob* have higher *Beta*; this is expected because *BSM-Prob* is an increasing function of stock price volatility, which includes both systematic and idiosyncratic components. Consistent with Panel A, firms with higher bankruptcy risk tend to have smaller market capitalization. The pattern of *BE/ME* is similar to that in Panel A, except that the decline of book-to-market from the second highest decile portfolio to the highest decile portfolio is much milder in Panel B. The pattern of *Ret1yr* in Panel B displays a striking difference in comparison to that in Panel A. Firms in the highest *BSM-Prob* decile do not have the lowest subsequent returns. Rather, the lowest two *BSM-Prob* deciles have the two lowest average subsequent returns. Moreover, firms in the second highest *BSM-Prob* decile have the highest subsequent returns. Combined, the results in Table 2.1 confirm Dichev's (1998) findings that *O-Score* is positively related to future returns. However, the results using *BSM-Prob* challenge his conclusion that market prices do not fully impound the implications of available bankruptcy information.

An interesting finding in Panel A of Table 2.1 is that the pattern of *Ret1yr* goes against the size effect, i.e., prior studies find that firms with a smaller market capitalization tend to have higher returns. Hence, I expect that the relationship between *O-Score* and *Ret1yr* will be more evident once firm size is controlled for. Another interesting finding is that the pattern of *Ret1yr* closely resembles that of *BE/ME*, which casts doubt on the existence of (any) relationship between *O-Score* and *Ret1yr*. In other words, it is not clear whether the pattern of subsequent returns reported in Panel A is evidence of the negative association of returns with Ohlson's (1980) bankruptcy risk measure or a manifestation of the book-to-market effect. Thus, controlling for the book-to-market effect as well as the size effect seems indispensable for the investigation of an association between Ohlson's bankruptcy risk measure and subsequent returns.

Table 2.2 presents the results of Fama-MacBeth regressions of subsequent returns on each measure of bankruptcy risk along with *Beta*, *ME*, and *BE/ME*. For each year from 1978 to 2007, I rank the independent variables into deciles and then convert them to scaled decile ranks from −0.5 to 0.5 by the following transformation: subtract 1, divide by 9 and subtract 0.5. After the conversion, the dependent variable is regressed on the scaled decile ranks of the independent variables. The mean of the coefficient estimates over the 30 sample years is reported along with the *t*-statistic, which is the mean of the coefficient estimates divided by the time-series standard error of the coefficient estimates. The use of decile ranks controls for outliers and non-linearity and facilitates interpretation of the economic magnitudes of the independent variables in relation to the dependent variable. An additional advantage of using decile ranks is that the coefficient on the independent variable represents the net return from a zero investment strategy of going long on the highest decile of the independent variable under consideration and short on the lowest decile while controlling for other variables. The scheme of scaled decile ranks has been commonly used in other studies that examine market inefficiency, especially studies that examine post-earnings-announcement drift (e.g., Mendenhall, 2004; Narayanamoorthy, 2006).

In Panel A of Table 2.2, the mean coefficient on the decile ranks of *O-Score* for the full sample is −6.06, with a *t*-statistic of −3.36. This suggests that one can obtain on average a 6.06 percent return (before transaction costs) by buying firms in the lowest decile of *O-Score* and shorting firms in the highest decile over the period from 1978 to 2007, after controlling for *Beta*, *ME*, and *BE/ME*. Using only firms with a December 31 year-end, I find that the mean coefficient on the decile ranks of *O-Score* is similar to that for the full sample and the return-predictive power of *O-Score* is not subsumed by the size and the book-to-market effects.

The regression analyses using subsamples confirm Griffin and Lemmon's (2002) findings that the negative relationship between Ohlson's (1980) bankruptcy risk measure and subsequent returns is primarily from firms with a small market capitalization and low book-to-market ratios (growth firms). The mean coefficients on the decile ranks of *O-Score* for the subsamples of small and growth firms are negative and significantly different from zero, whereas those for the subsamples of large and value firms are not reliably different from zero.

In both Panels A and B of Table 2.2 (as well as throughout this study), the positive and significant intercepts represent the average annual raw returns from investing in firms with median values of each variable over the sample period. The mean coefficients on *Beta* are insignificant, which is expected given that the CAPM has failed to describe average realized stock returns since the early 1960s (Campbell and Vuolteenaho, 2004). Interestingly, the mean coefficients on *ME* are also insignificant, but this is consistent with other studies that also document that the size effect is weak to nonexistent in the 1980s and 1990s (e.g., Fama and French, 1992). The mean coefficients on *BE/ME* are economically and statistically different from zero, except the one for the subsample of value firms, which is likely due to the relatively small cross-sectional variations in *BE/ME* in the subsample of value firms.

In Panel B of Table 2.2, none of the mean coefficients on the decile ranks of *BSM-Prob* are significantly different from zero. Recall that *BSM-Prob* has been shown to provide a better prediction of bankruptcy than *O-Score*. If market prices do not fully incorporate the implications of bankruptcy information, why is it the case that *BSM-Prob* is unrelated to subsequent returns?[13] Moreover, Griffin and Lemmon's (2002) explanation of Dichev's (1998) findings does not seem to apply when *BSM-Prob* is the proxy of bankruptcy risk: the mean coefficients on the decile ranks of *BSM-Prob* for the subsamples of small and growth firms are not reliably different from

[13] Note that *BSM-Prob* is a measure based on market prices. If the market is not adequately anticipating bankruptcy probabilities, then it is possible for *BSM-Prob* not to show the bankruptcy anomaly even if the anomaly exists. I can not confirm or refute this possibility.

zero. Taken as a whole, Table 2.2 suggests that Dichev's conclusion is potentially idiosyncratic to Ohlson's (1980) measure of bankruptcy risk. This raises the question: why is *O-Score* negatively associated with subsequent returns? This is the question I explore in the next section.

2.3 Empirical analysis II: What is inside the magic of O-Score?

The measure of bankruptcy risk proposed by Ohlson (1980) is a score – a linear combination of accounting variables. As demonstrated in Appendix B, when there is an association between a variable and a score, it must be true that one or more components of the score have nontrivial associations with the variable, holding all other components constant. Thus, the association between returns and Ohlson's measure of bankruptcy risk ensures that one or more components of O-Score are associated with subsequent returns.

2.3.1 Results

To examine which components of *O-Score* are responsible for the association of returns with the score, I run Fama-MacBeth regressions of subsequent returns on the components of *O-Score* and report the results in Table 2.3. This table is similar to Panel A of Table 2.2, except that I use the components of *O-Score* rather than *O-Score* itself.

The regression results in Table 2.3 reveal that *FFO/TL* is the only component that is consistently associated with subsequent returns. Other components are also associated with subsequent returns but the associations are only observed for a single subset of firms. By contrast, the association of returns with *FFO/TL* is economically and statistically different from zero in each and every case. For example, the mean coefficient on the decile ranks of *FFO/TL* for the full sample is 6.87, with a *t*-statistic of 3.21. This suggests that one can obtain on average a 6.87 percent return (before transaction costs) by buying firms in the highest decile of *FFO/TL* and shorting firms in the lowest decile over the period from 1978 to 2007, after controlling for the other components of *O-Score*, *Beta*, *ME*, and *BE/ME*. Note that the mean coefficients on the decile ranks of *O-Score* for the subsamples of large and value firms are not reliably different from zero in Panel A of Table 2.2, whereas the corresponding mean coefficients on the decile ranks of *FFO/TL* in Table 2.3 are significantly different from zero. While an association of a variable with a score implies an association of the variable with at least one component of the score, the reverse does not follow.

To corroborate the results in Table 2.3, I assess the profitability of a trading strategy using *FFO/TL*. At the end of June of each year from 1978 to 2007, I construct a zero-investment trading strategy based on buying an equal-weighted (value-weighted) portfolio of firms with the highest decile ranks of *FFO/TL* and shorting an equal-weighted (value-weighted) portfolio of firms with the lowest decile ranks and report the annual returns in Panel A (Panel B) of Table

2.4.[14] Consistent with Table 2.3, firms with high *FFO/TL* earn higher returns than firms with low *FFO/TL*.[15] The equal-weighted as well as value-weighted trading strategy generates economically sizable profits over my sample period. The mean and median annual return on the equal-weighted (value-weighted) trading strategy is 7.53 and 12.47 percent (13.03 and 15.25 percent), respectively. In addition, the profitability of the strategy is consistent; firms with high *FFO/TL* outperform firms with low *FFO/TL* in 21 (22) out of 30 years in Panel A (Panel B).

The results presented so far suggest that the *FFO/TL* component of *O-Score* is associated with returns. Next, I examine whether the return-predictive power of *O-Score* disappears once *FFO/TL* is controlled for. Table 2.5 presents the results of Fama-MacBeth regressions of *Ret1yr* on both *O-Score* and *FFO/TL* (along with *Beta*, *ME*, and *BE/ME*). The mean coefficients on the decile ranks of *O-Score* are no longer negative. Thus, the results in Table 2.5, in conjunction with those in Panel A of Table 2.2, suggest that the association of returns with *O-Score* is driven by the association of returns with its component, *FFO/TL*. Note that the mean coefficients on the decile ranks of *FFO/TL* in Table 2.5 are larger than those in Table 2.3 but they are statistically

---

[14] I include value-weighted portfolio analysis because Fama (1998) finds that when value-weighted returns are examined, evidence of stock market inefficiency shrinks substantially and typically becomes statistically unreliable.

[15] The trading strategy in Table 2.4 (as well as in Table 2.8) earned large negative net returns in 1999. This period is believed to be a time in which mispricing was prevalent (see Ritter and Welch, 2002; Lamont and Thaler, 2003).

less significant, which is expected given the variance inflation resulting from the high correlation between *O-Score* and *FFO/TL*.[16]

2.4 Empirical analysis III: What is the source of the predictive power of FFO/TL?

Results presented to this point show that the return-predictive power of O-Score found in Dichev (1998) is due to the association of returns with its component, FFO/TL. Prompted by a close relationship between funds from operations (FFO) and cash flows from operations (CFO), I conjecture that the rankings of funds from operations mimic those of cash flows from operations. The relation between funds from operations and cash flows from operations, along with the accrual anomaly, is discussed in Section 2.4.1. The description of the sample and the definition of the variables are provided in Section 2.4.2, followed by empirical results in Section 2.4.3 and the results of sensitivity tests are reported in Section 2.4.4.

2.4.1 Funds from operations, cash flows from operations, and accrual anomaly

FFO was one of the disclosure requirements in the statement of changes in financial position (SCFP) before the statement of cash flows became mandatory.[17] Most firms prepared SCFP on a

---

[16] The Pearson (Spearman) correlation coefficient is –0.74 (–0.85).

[17] In 1963, the Accounting Principle Board (APB) issued APB Opinion No. 3, *The Statement of Source and Application of Funds* (AICPA, 1963). This standard encouraged firms to present a funds statement, a predecessor of SCFP, as an integral part of a company's financial statements.

working capital basis, meaning that FFO represented working capital generated from operations. In this case FFO is essentially CFO before the adjustments of non-cash working capital in the indirect method of presenting the statement of cash flows. Some firms reported SCFP on a cash basis. In this case FFO is the same as CFO.

Ohlson's (1980) models of bankruptcy prediction use FFO divided by total liabilities. The idea is to capture the degree to which the internally generated working capital of a firm is available to satisfy its debt obligations. As firms with large liabilities tend to have large recorded assets, the cross-sectional rankings of FFO/TL are likely to follow those of CFO/ATA, which has been found to be associated with subsequent returns in Sloan (1996) and Houge and Loughran (2000).

Sloan (1996) finds that firms with higher accruals tend to have lower subsequent returns, i.e., accruals have a negative association with subsequent returns. Sloan conjectures that investors fixate on current earnings and fail to fully price the different information contained in the accruals and cash flows components of earnings. In regard to testing his hypotheses, Sloan does not use the magnitude of accruals for his tests. Rather, he uses relative accruals, i.e., accruals deflated by average total assets to control for size. While the trading strategy in Sloan is based on accruals, he also documents in his footnote 15 (p. 306) that the negative association between accruals and cash flows ensures that the trading strategy using CFO/ATA produces positive returns of a similar magnitude; this is expected as earnings are the sum of accruals and cash flows. Indeed, a subsequent study by Houge and Loughran (2000) shows that CFO/ATA has a positive association with subsequent returns with a dataset spanning 1963 to 1993, confirming Sloan's findings that investors overestimate the persistence of accruals relative to that of cash flows.

2.4.2 Sample formation and variable measurement

I follow Sloan (1996) and Houge and Loughran (2000) with regard to variable definitions. I define the variable *CFO/ATA* as cash flows from operations divided by average total assets. Cash flows from operations are operating income after depreciation minus accruals.[18] Accruals are defined as the change in non-cash current assets minus the change in current liabilities, excluding the current portion of long-term debt and taxes payable, minus depreciation expense, i.e., (ΔCA − ΔCash) − (ΔCL − ΔSTD − ΔTP) − Depreciation. I use the financial statements of fiscal years ending in calendar year $t-1$ to calculate CFO/ATA_t so that the variable is estimated on an *ex ante* basis. The empirical analysis in this section begins with the sample used in the previous sections. After restricting firm-years to those with sufficient data to compute the variables described above, my final sample is 74,364 firm-years representing 8,594 firms.

---

[18] Note that cash flows from operations could be derived from reported funds from operations for observations on pre-SFAS No. 95 regime, but it is standard practice to adopt Sloan's definitions of accruals (and cash flows from operations, if applicable) among studies that examine the accrual anomaly.

2.4.3 Results

To demonstrate the relationship between *FFO/TL* and *CFO/ATA*, I form portfolios according to the rankings of *FFO/TL* at the end of June of each year from 1978 to 2007. The average values of the two variables along with other variables are reported in Table 2.6. As can be seen, the pattern of *Ret1yr* confirms that *FFO/TL* has return-predictive power. Although the pattern is not monotonic, firms in the two lowest *FFO/TL* deciles have the lowest subsequent returns. The close relationship between *FFO/TL* and *CFO/ATA* is demonstrated in the pattern of *CFO/ATA*. The monotonically increasing pattern of *CFO/ATA* demonstrates that the rankings of *FFO/TL* indeed mimic those of *CFO/ATA*. *BE/ME* does not display any meaningful pattern, which in turn suggests that the association of returns with *FFO/TL* is not driven by the book-to-market effect.

Table 2.7 presents the results of Fama-MacBeth regressions of *Ret1yr* on C*FO/ATA* (along with *Beta*, *ME*, and *BE/ME*). The mean coefficient on the decile ranks of *CFO/ATA* for the full sample is 11.04, with a *t*-statistic of 4.27. This suggests that one can obtain on average an 11.04 percent return (before transaction costs) by buying firms in the highest decile of *CFO/ATA* and shorting firms in the lowest decile over the period from 1978 to 2007, after controlling for *Beta*, *ME*, and *BE/ME*. Just as *FFO/TL* is found to be positively associated with returns within each subset of firms in Table 2.3, there is a positive association between *CFO/ATA* and returns in Table 2.7, with the mean coefficient on the decile ranks of *CFO/ATA* ranging from 7.10 to 11.39. In summary, Table 2.7 demonstrates the positive association of returns with *CFO/ATA*.

To corroborate the results in Table 2.7, I assess the profitability of a trading strategy using *CFO/ATA*. At the end of June of each year from 1978 to 2007, I construct a zero-investment trading strategy based on buying an equal-weighted (value-weighted) portfolio of firms with the highest decile ranks of *CFO/ATA* and shorting an equal-weighted (value-weighted) portfolio of firms with the lowest decile ranks and report annual return in Panel A (Panel B) of Table 2.8. The returns to trading strategies reported in Table 8 is comparable to those based on *FFO/TL* in Table 2.4. The mean and median annual return on the equal-weighted (value-weighted) trading strategy is 10.99 and 14.97 percent (12.17 and 21.16 percent), respectively. As in Table 4, the profitability of the strategy is consistent; firms with high *CFO/ATA* outperform firms with low *CFO/ATA* in 24 (20) out of 30 years in Panel A (Panel B) of Table 2.8.

To enhance the comparability between Tables 2.4 and 2.8, annual net returns from the zero-investment trading strategy using *FFO/TL* and *CFO/ATA* are plotted side by side in Figure 2.1. The figure reveals strong similarities between the results from the two different strategies. The signs of net returns from the two strategies agree in 27 (26) out of 30 years in Panel A using equal-weighted portfolios (Panel B using value-weighted portfolios). Moreover, when the strategy using *CFO/ATA* fails to produce positive net returns, the strategy using *FFO/TL* also fails to produce positive net returns, with few exceptions. The results presented so far suggest that the association of returns with *FFO/TL* is related to the association of returns with *CFO/ATA*. Next, I examine whether the return-predictive power of *FFO/TL* disappears once *CFO/ATA* is controlled for.

Table 2.9 presents the results of Fama-MacBeth regressions of *Ret1yr* on *FFO/TL* before and after controlling for *CFO/ATA* in Panels A and B, respectively. Panel A shows that *FFO/TL* predicts subsequent returns, but Panel B demonstrates that the association of returns with *FFO/TL* diminishes once *CFO/ATA* is controlled for. While the mean coefficients on the decile ranks of *FFO/TL* are still statistically different from zero in Panel B, their economic significance is reduced by roughly half when compared to Panel A. For example, the mean coefficient on the decile ranks of *FFO/TL* for the full sample is 8.07 with a *t*-statistic of 4.38 in Panel A, whereas the corresponding mean coefficient in Panel B is 3.51 with a *t*-statistic of 2.25. Note that the variable *CFO* is derived from the accounting information in the balance sheet and the income statement, whereas the variable *FFO* is derived from the statement of changes in financial position (the statement of cash flows) before (after) SFAS No. 95 becomes effective. When I limit my analysis to post-SFAS No. 95 observations and derive both variables from the statement of cash flows, the decile ranks of FFO/TL are no longer associated with returns (see Section 2.4.4).

In Section 2.3, I showed that the association of returns with *O-Score* is due to the association of returns with *FFO/TL*. Hence, if the association of returns with *FFO/TL* is driven by the association of returns with *CFO/ATA*, I expect the return-predictive power of *O-Score* to diminish or disappear once *CFO/ATA* is controlled for. Table 2.10 shows that once *CFO/ATA* is controlled for, the mean coefficients on the decile ranks of *O-Score* are no longer statistically and economically different from zero, except for the subsample of small firms. Throughout Tables 2.9 and 2.10, the mean coefficients on the decile ranks of *CFO/ATA* are economically and statistically different from zero, suggesting that the return-predictive power of *CFO/ATA* is not subsumed by either *FFO/TL* or *O-Score*. Thus, the results presented so far support my prediction that the association of returns with *FFO/TL* (and with *O-Score*) is largely due to the association of returns with *CFO/ATA*.

To corroborate the regression analysis, I conduct a two-way sort according to an observation's quintile ranking of *FFO/TL* (*O-Score*) and *CFO/ATA* at the end of June of each year from 1978 to 2007. Each Panel in Table 2.11 reports average annual returns for each of 25 portfolios. It also reports the average number of firms over the sample years for each portfolio in bracket and the average of the difference in *Ret1yr* between the extreme quintiles, followed by a time-series *t*-statistic in parenthesis. Note that with regard to *O-Score*, I report *Ret1yr* of the lowest quintile minus that of the highest quintile because firms with lower *O-Score* are found to have higher subsequent returns.

The concentration of firms in the main diagonal cells in Panel A is expected given the positive correlation between *FFO/TL* and *CFO/ATA*. As shown in the bottom two rows of Panel A, within each *CFO/ATA* quintile column, firms with higher *FFO/TL* do not necessarily have higher subsequent returns. However, as shown in the last column, for four out of five *FFO/TL* quintile rows, firms with higher *CFO/ATA* earn higher subsequent returns. These results in Panel

A support the findings in Table 2.9 that the association of returns with *FFO/TL* is due to the association of returns with *CFO/ATA*.

In Panel B, the concentration of firms in the minor diagonal cells demonstrates a negative correlation between *O-Score* and *CFO/ATA*. The pattern in each column, with the exception of the first, suggests that firms with higher *O-Score* do not earn lower subsequent returns, once *CFO/ATA* is controlled for. Recall that the negative association of returns with *O-Score* is primarily driven by the poor performance of firms with high *O-Score* (see Panel A of Table 2.1). An examination of firms in the highest *O-Score* quintile row unveils striking evidence that only the quintile of firms with low *CFO/ATA* earns lower than average returns (4.63 percent). Taken as a whole, the evidence reported in Panel B confirms the findings that the negative association of returns with Ohlson's (1980) measure of bankruptcy risk is a manifestation of Sloan's (1996) accrual anomaly.

2.4.4 Sensitivity tests

The measure of cash flows utilized to this point is the difference between operating income after depreciation minus accruals derived from the balance sheet and income statement. This method of estimating the cash flow involves measurement error relative to measuring accruals directly from the statement of cash flows, as noted in Hribar and Collins (2002). Thus, as a sensitivity test, I limit the analysis to firm-years in the post-SFAS No. 95 regime and then repeat the analysis using a cash flow measure from the statement of cash flows.

Specifically, I use cash flows from continuing operations, which are cash flows from operations minus the cash portion of discontinued operations and extraordinary items, as suggested by Hribar and Collins (2002). Note that in the main analysis, the variable *CFO* is derived from the accounting information in the balance sheet and the income statement, whereas the variable *FFO* is derived from the statement of changes in financial position (the statement of cash flows) before (after) SFAS No. 95 became effective. Thus, the analysis in this section effectively eliminates any possible inconsistency in the calculation of *FFO* and *CFO* as both variables are derived from the statement of cash flows. Furthermore, using observations in the post-SFAS No. 95 regime ensures cross-sectional comparability in the calculation of *FFO* because I no longer resort to reported funds from operations. However, these benefits come at the cost of reduction in power in a statistical test because the sample includes only observations in the post-SFAS No. 95 regime.

Table 2.12 presents the replication of Tables 2.7, 2.9, and 2.10 using only the post-SFAS No. 95 observations. The sample consists of 47,948 firm-years (6,309 firms), spanning 19 years from 1989 to 2007. For brevity, I only report the results of the variables of interest. Throughout Table 2.12, the mean coefficients on the decile ranks of *CFO/ATA* are more or less of a similar magnitude to the corresponding mean coefficients in the matching tables. However, they become less significant because the analysis only employs 19 time-series. Consistent with Table 2.7, Panel A demonstrates a positive association between subsequent returns and *CFO/ATA*. In Table

2.9, the coefficients on *FFO/TL* become smaller but still statistically significantly different from zero when *CFO/ATA* is controlled for. By contrast, Panel B of Table 2.12 shows that *FFO/TL* is no longer associated with subsequent returns once *CFO/ATA* is controlled. This is so because the analysis in Panel B effectively eliminates inconsistency in the calculation of *FFO* and *CFO* (i.e., both variables are derived from the statement of cash flows). Panel C confirms that a negative association of returns with *O-Score* disappears once *CFO/ATA* is controlled for.

The results provided so far support that Dichev's (1998) bankruptcy risk anomaly is essentially Sloan's (1996) accrual anomaly in disguise and yet rely on cash flows to proxy the accrual anomaly in lieu of the conventional accrual measure. As noted in Sloan, the two components of earnings are highly correlated, but their correlation is unlikely to be perfect. Thus, the natural question that follows is whether the negative association between O-Score and returns would disappear when I employ Sloan's original accrual measure.[19]

Panel A of Table 2.13 presents the results of Fama-MacBeth regressions of *Ret1yr* on *O-Score* and *ACC/ATA* along with *Beta*, *ME*, and *BE/ME*, where *ACC/ATA* represents Sloan's (1996) original accrual measure. To be consistent with the main analysis, I convert independent variables into scaled decile ranks. As in Table 2.12, Table 2.13 only reports the results of the variables of interest. The most striking observation from Panel A is that, contrary to the results in

[19] I want to express my gratitude to an anonymous referee for raising this question.

the main analysis, the negative association of *O-Score* with returns does not disappear. Thus, the evidence provided in Panel A seems to contradict the evidence provided in the main analysis. However, that is not necessarily the case. When two highly correlated independent variables are regressed, their explanatory power and the significance of their coefficients are divided up between the two. Since the ability of *O-Score* to predict returns is not distinct from the ability of *CFO/ATA* to predict returns, *O-Score* will "dominate" *ACC/ATA* to the extent that *CFO/ATA* better explains the variation in returns than *ACC/ATA*. Moreover, this scenario provides good explanation for the marginal association between *ACC/ATA* and returns.

To investigate which proxy of Sloan's (1996) accrual anomaly better predicts returns, I run Fama-MacBeth regressions of *Ret1yr* on both *ACC/ATA* and *CFO/ATA* along with *Beta*, *ME*, and *BE/ME* and present the results in Panel B of Table 2.13. Once cash flows are controlled for, Sloan's original accrual measure is no longer associated with returns, which, in turn, explains the results in Panel A. Panel C reports the association of *O-Score* after controlling for *ACC/ATA* and *CFO/ATA*. As expected, *CFO/ATA* is the only variable that is associated with future returns.

2.5 Conclusion

Using Ohlson's (1980) measure of bankruptcy risk, Dichev (1998) documents that firms with high bankruptcy risk earn lower than average returns in subsequent years. This anomaly suggests that market prices do not fully impound the implications of available bankruptcy information. This study challenges that conclusion.

The first empirical analysis in this study investigates the robustness of the bankruptcy risk anomaly using O-Score and BSM-Prob within identical research settings. With a dataset spanning 1978 to 2007, I find a reliable negative association between O-Score and returns, but fail to find any association between BSM-Prob and returns, suggesting that the negative association between bankruptcy risk and returns is construct-dependent. I then examine the individual components of O-Score and find that the funds from operations component (FFO/TL) is the only one that is associated with future returns. I show that the association of returns with this component is responsible for the association of O-Score and returns. In light of these results, the final empirical analysis investigates the source of the return-predictive power of FFO/TL. FFO/TL and CFO/ATA are closely related and thus, I conjecture and find that the cross-sectional rankings of FFO/TL follow those of CFO/ATA, an alternative implementation vehicle of Sloan's (1996) accrual anomaly. Taken as a whole, the results of this study suggest that the negative association between Ohlson's (1980) measure of bankruptcy risk and subsequent returns is a manifestation of investors' under (over)-pricing of cash flows (accruals) component of earnings, i.e., Sloan's accrual anomaly.

This study is the first attempt to relate the bankruptcy risk anomaly to the accrual anomaly. It clearly demonstrates that the two seemingly unrelated stock market anomalies share the same underlying cause. Given the substantial number of anomalies in the literature, it is possible, and even likely, that there are more such commonalities among other anomalies. Therefore, researchers should carefully consider alternative explanations, including existing anomalies, before claiming the discovery of a new anomaly.

Table 2.1: Characteristics of O-Score [BSM-Prob] decile portfolios

Panel A: Characteristics of O-Score decile portfolios

	Portfolio O-Score ranking									
	lowest	2	3	4	5	6	7	8	9	highest
O-Score	−8.63	−6.56	−5.76	−5.16	−4.63	−4.09	−3.52	−2.78	−1.66	2.67
O-Prob	0.04	0.15	0.33	0.60	1.02	1.75	3.14	6.55	18.46	67.11
Beta	1.24	1.14	1.11	1.11	1.09	1.08	1.10	1.13	1.17	1.34
ME	5.43	5.56	5.55	5.33	5.01	4.63	4.21	3.72	3.14	2.82
BE/ME	0.57	0.70	0.75	0.81	0.87	0.93	0.99	1.02	0.92	0.04
Ret1yr	14.49	16.23	16.58	16.97	16.62	16.32	17.31	17.81	12.02	6.16

Panel B: Characteristics of BSM-Prob decile portfolios

	Portfolio BSM-Prob ranking									
	lowest	2	3	4	5	6	7	8	9	highest
BSM-Prob	0.00	0.00	0.00	0.00	0.02	0.13	0.52	1.87	6.27	28.31
Beta	0.90	1.05	1.10	1.15	1.16	1.21	1.22	1.23	1.26	1.26
ME	6.05	5.83	5.50	5.19	4.81	4.50	4.12	3.70	3.17	2.46
BE/ME	0.41	0.55	0.62	0.66	0.72	0.79	0.89	0.98	1.08	0.90
Ret1yr	13.11	13.00	14.64	14.53	14.83	15.44	16.06	16.99	18.32	13.44

Note to Table 2.1:

Panel A (Panel B) of this table presents mean values of characteristics for ten portfolios formed annually at the end of June of year *t* from 1978 to 2007 by assigning firms to deciles based on the magnitude of $O\text{-}Score_t$ ($BSM\text{-}Prob_t$). The sample consists of 74,737 firm-years representing 8,616 firms. *O-Score* is a measure of bankruptcy risk from Model 1 in Ohlson (1980) and higher values of *O-Score* signify higher probability of bankruptcy. *O-Prob* is a probability form of *O-Score* derived using the logistic cumulative distribution function. *BSM-Prob* is a probability of bankruptcy that is from Hillegeist et al. (2004). $Beta_t$ is estimated using the Capital Asset Pricing Model (CAPM) for each firm using monthly returns for at least 24 of the 60 months preceding July of year *t*. ME_t is log of a firm's market equity in millions at the end of June of year *t*. BE/ME_t is book common equity for the fiscal year ending in calendar year $t-1$, divided by market equity at the end of December $t-1$. $Ret1yr_t$ is the buy-and-hold returns (in percent) over one-year period beginning July of year *t*.

Table 2.2: Association of returns with O-Score [BSM-Prob]
Dependent variable = Ret1yr

	All Firms (N= 74,737)	December Firms (N= 38,859)	Small Firms (N=49,915)	Large Firms (N=24,822)	Growth Firms (N=43,866)	Value Firms (N=30,871)
Panel A: Association of returns with O-Score						
Intercept	15.05***	14.94***	15.54***	14.36***	12.01***	19.33***
	(3.55)	(3.74)	(3.19)	(4.33)	(2.86)	(4.30)
O-Score$^{\mathbf{D}}$	−6.06***	−5.79**	−7.86***	−2.20	−5.76***	−3.01
	(−3.36)	(−2.53)	(−3.83)	(−0.96)	(−2.79)	(−1.30)
Beta$^{\mathbf{D}}$	−1.64	−1.40	0.01	−4.51	−2.86	0.28
	(−0.34)	(−0.30)	(0.00)	(−0.94)	(−0.58)	(0.06)
ME$^{\mathbf{D}}$	−2.43	−1.83	−3.64	1.53	0.40	−6.40
	(−0.53)	(−0.40)	(−1.04)	(0.66)	(0.09)	(−1.42)
BE/ME$^{\mathbf{D}}$	13.12***	13.73***	14.45***	5.91**	11.56***	1.71
	(3.94)	(3.92)	(4.12)	(2.05)	(4.67)	(0.63)
Avg. Adj. R^2 (%)	4.04	4.52	3.28	6.30	4.73	2.13
Panel B: Association of returns with BSM-Prob						
Intercept	15.06***	14.94***	15.54***	14.36***	12.00***	19.33***
	(3.55)	(3.74)	(3.19)	(4.34)	(2.86)	(4.30)
BSM-Prob$^{\mathbf{D}}$	0.78	1.90	0.18	0.41	−0.27	2.86
	(0.34)	(0.69)	(0.07)	(0.22)	(−0.13)	(0.95)
Beta$^{\mathbf{D}}$	−2.04	−2.23	−0.60	−4.25	−2.68	−0.44
	(−0.46)	(−0.51)	(−0.13)	(−0.91)	(−0.55)	(−0.11)
ME$^{\mathbf{D}}$	1.31	2.52	−0.32	1.92	3.31	−3.73
	(0.31)	(0.57)	(−0.09)	(0.83)	(0.75)	(−0.85)
BE/ME$^{\mathbf{D}}$	14.25***	14.40***	16.41***	6.09*	13.08***	1.02
	(4.29)	(3.94)	(4.60)	(2.01)	(5.34)	(0.36)
Avg. Adj. R^2 (%)	4.07	4.57	3.21	6.15	4.71	2.27

Note to Table 2.2

This table presents the results of Fama-MacBeth regressions using 74,737 firm-years (8,616 firms), spanning 30 years from 1978 to 2007. The dependent variable is regressed on scaled decile ranks, ranging from −0.5 to 0.5, of the independent variables each year. The regression coefficient reported here is the mean of the coefficient estimates over 30 sample years and *t*-statistic reported in parenthesis is the mean of the coefficient estimates divided by time-series standard error of the coefficient estimates. Variable definitions are in Table 2.1. Superscript D ($^{\mathbf{D}}$) denotes decile ranks. December Firms are firms with a

December 31 year-end. Small (Large) Firms include those firms with market capitalization smaller (larger) than the first quintile of NYSE size distribution. Growth (Value) Firms include firms whose book-to-market is smaller (larger) than the median of NYSE book-to-market distribution. *, **, *** denote two-tail significance at the 10%, 5%, and 1% level, respectively.

Table 2.3: Association of returns with the components of O-Score
Dependent variable = Ret1yr

	All Firms (N=74,737)	December Firms (N=38,859)	Small Firms (N=49,915)	Large Firms (N=24,822)	Growth Firms (N=43,866)	Value Firms (N=30,871)
Intercept	15.15***	15.06***	15.61***	14.26***	12.27***	19.37***
	(3.67)	(3.88)	(3.27)	(4.33)	(2.94)	(4.46)
Components of O-Score						
Size^D	–1.89	–5.01	–4.02	–5.87	–1.07	–23.29***
	(–0.37)	(–0.86)	(–1.03)	(–0.98)	(–0.17)	(–3.09)
TL/TA^D	1.75	2.63	1.44	5.45*	3.64	3.85
	(0.47)	(0.72)	(0.33)	(1.86)	(0.88)	(0.92)
WC/TA^D	5.78	7.86	7.01	2.84	12.54**	1.34
	(1.26)	(1.44)	(1.44)	(0.90)	(2.23)	(0.27)
CL/CA^D	3.19	4.46	2.62	2.48	8.81**	–0.22
	(1.10)	(1.28)	(0.81)	(0.99)	(2.30)	(–0.07)
OENEG	–3.63	0.54	–4.11	–0.70	–1.90	N/A
	(–1.22)	(0.15)	(–1.12)	(–0.13)	(–0.63)	
NI/TA^D	1.23	–0.99	0.12	–0.09	–0.17	–6.57
	(0.41)	(–0.30)	(0.03)	(–0.03)	(–0.04)	(–1.69)
FFO/TL^D	6.87***	8.30***	8.91***	7.50***	8.82***	7.65***
	(3.21)	(3.18)	(3.14)	(3.09)	(3.23)	(2.83)
INTWONEG	–1.06	–1.79	–0.89	–0.12	–1.81	–1.18
	(–0.73)	(–1.07)	(–0.56)	(–0.05)	(–1.25)	(–0.66)
CHIN^D	1.01	1.62	0.82	2.02	0.12	4.45
	(0.76)	(1.22)	(0.47)	(1.57)	(0.10)	(1.52)
Beta^D	–0.96	–0.57	0.82	–2.84	–1.80	1.36
	(–0.27)	(–0.17)	(0.22)	(–0.81)	(–0.53)	(0.34)
ME^D	0.17	4.24	–0.51	6.40	2.27	16.80**
	(0.03)	(0.69)	(–0.11)	(1.28)	(0.35)	(2.53)
BE/ME^D	14.23***	15.58***	16.35***	10.15***	11.87***	5.64**
	(5.14)	(5.26)	(5.36)	(3.45)	(4.63)	(2.19)
Avg. Adj. R^2 (%)	5.81	6.81	4.85	9.39	6.81	4.10

Note to Table 2.3:

This table presents the results of Fama-MacBeth regressions using 74,737 firm-years (8,616 firms), spanning 30 years from 1978 to 2007. The dependent variable is regressed on scaled decile ranks, ranging from –0.5 to 0.5, of the independent variables, with the exception of the indicator variables, *OENEG* and *INTWONEG*, each year. The regression coefficient reported here is the mean of the coefficient estimates over 30 sample years and *t*-statistic reported in parenthesis is the mean of the coefficient estimates divided by time-series standard error of the coefficient estimates. Definitions of the components of *O-Score* are as follows: *Size* is the log(total assets/GDP price level index); *TL/TA* is total liabilities divided by total assets; *CL/CA* is current liabilities divided by current assets; *NI/TA* is net income divided by total assets; *FFO/TL* is funds from operations divided by total liabilities; *INTWONEG* is an indicator variable equal to one if cumulative net income over the previous two years is negative, and zero otherwise; *OENEG* is an indicator variable equal to one if owner's equity is negative, and zero otherwise; *CHIN* = $(NI_t - NI_{t-1})/(|NI_t| + |NI_{t-1}|)$ is the scaled change in net income. Refer to Table 2.1 for the definitions of *Ret1yr*, *Beta*, *ME*, and *BE/ME*. Superscript D (^{D}) denotes decile ranks. December Firms are firms with a December 31 year-end. Small (Large) Firms include those firms with market capitalization smaller (larger) than the first quintile of NYSE size distribution. Growth (Value) Firms include firms whose book-to-market is smaller (larger) than the median of NYSE book-to-market distribution. *, **, *** denote two-tail significance at the 10%, 5%, and 1% level, respectively.

Table 2.4: Returns to zero-investment trading strategies using FFO/TL

	Panel A: Equal-weighted portfolio			Panel B: Value-weighted portfolio		
Year	FFO/TL Highest Decile Ret	FFO/TL Lowest Decile Ret	Net Ret	FFO/TL Highest Decile Ret	FFO/TL Lowest Decile Ret	Net Ret
1978	17.60	11.00	6.61	13.02	11.63	1.39
1979	22.95	7.45	15.50	−0.64	13.20	−13.84
1980	41.60	41.74	−0.13	21.29	26.32	−5.03
1981	−18.00	−26.38	8.38	−4.65	−29.63	24.98
1982	80.58	111.66	−31.08	73.83	106.70	−32.87
1983	−16.14	−30.85	14.71	−12.89	−32.05	19.17
1984	9.85	−20.04	29.89	20.65	−19.58	40.23
1985	25.70	9.81	15.89	31.35	15.61	15.74
1986	7.19	−3.04	10.23	15.84	−7.59	23.43
1987	−11.28	−29.60	18.32	−17.30	−32.36	15.06
1988	7.47	−8.75	16.22	1.61	−12.37	13.98
1989	16.01	−8.34	24.35	15.27	4.86	10.41
1990	6.22	10.23	−4.01	9.23	10.45	−1.22
1991	14.36	28.65	−14.29	10.79	5.65	5.14
1992	19.56	23.26	−3.70	15.77	−18.51	34.28
1993	1.47	−17.26	18.73	−7.07	−35.51	28.44
1994	39.74	23.18	16.56	58.45	15.75	42.70
1995	19.64	48.77	−29.14	18.21	35.21	−17.00
1996	7.05	−32.70	39.75	33.77	−25.87	59.64
1997	11.03	−15.24	26.27	27.88	−15.42	43.30
1998	4.34	0.57	3.77	49.35	4.48	44.87
1999	57.88	138.39	−80.51	98.89	124.52	−25.63
2000	5.81	−25.81	31.62	−51.20	−27.31	−23.89
2001	−4.79	−40.17	35.38	−33.00	−56.54	23.54
2002	9.49	20.88	−11.39	13.35	9.14	4.21
2003	36.26	41.21	−4.94	34.75	32.77	1.98
2004	14.09	−19.41	33.50	−1.57	−24.88	23.31
2005	14.80	5.87	8.93	1.17	2.40	−1.24
2006	14.03	3.88	10.15	20.00	−0.31	20.30
2007	−10.20	−30.46	20.26	−6.23	−21.67	15.45
Mean			7.53^{*}			13.03^{***}
(t)			(1.71)			(3.17)
Median			12.47^{**}			15.25^{***}
(signed rank)			(119.5)			(134.5)
No. of positive			$21/30^{**}$			$22/30^{**}$
(Z)			(2.19)			(2.56)

Note to Table 2.4:

This table presents returns to zero-investment trading strategies using funds from operations divided by total liabilities (*FFO/TL*) from 74,737 firm-years (8,616 firms), spanning 30 years from 1978 to 2007. At the end of June of each year t, firms are assigned into deciles according to their rankings of FFO/TL_t. *Ret* is the return on each portfolio (in percent) over one-year period beginning July of year *t*. *Mean* at the bottom of the *Net Ret* column is the average of the *Net Ret* figures. *, **, *** denote two-tail significance at the 10%, 5%, and 1% level, respectively.

Table 2.5: Association of returns with O-Score after controlling for FFO/TL
Dependent variable = Ret1yr

	All Firms (N=74,737)	December Firms (N=38,859)	Small Firms (N=49,915)	Large Firms (N=24,822)	Growth Firms (N=43,866)	Value Firms (N=30,871)
Intercept	15.05***	14.94***	15.54***	14.35***	12.01***	19.33***
	(3.54)	(3.74)	(3.19)	(4.33)	(2.86)	(4.30)
O-Score^D	3.64	5.46	0.68	6.50*	9.18	2.98
	(0.74)	(0.96)	(0.13)	(1.76)	(1.56)	(0.75)
FFO/TL^D	10.91**	12.73**	10.25	10.53***	15.61***	7.64*
	(2.26)	(2.39)	(1.65)	(3.51)	(2.89)	(1.72)
Beta^D	−0.77	−0.12	0.99	−3.48	−1.60	0.89
	(−0.18)	(−0.03)	(0.23)	(−0.75)	(−0.36)	(0.20)
ME^D	−0.66	−0.03	−2.88	3.13	2.97	−5.07
	(−0.14)	(−0.01)	(−0.78)	(1.29)	(0.65)	(−1.06)
BE/ME^D	14.98***	15.52***	15.94***	8.40**	13.17***	2.85
	(4.16)	(4.11)	(4.12)	(2.66)	(5.09)	(0.97)
Avg. Adj. R^2 (%)	4.48	5.08	3.73	6.88	5.25	2.63

Note to Table 2.5:

This table presents the results of Fama-MacBeth regressions using 74,737 firm-years (8,616 firms), spanning 30 years from 1978 to 2007. The dependent variable is regressed on scaled decile ranks, ranging from −0.5 to 0.5, of the independent variables each year. The regression coefficient reported here is the mean of the coefficient estimates over 30 sample years and *t*-statistic reported in parenthesis is the mean of the coefficient estimates divided by time-series standard error of the coefficient estimates. Variable definitions are in Table 2.1. Superscript D (^{D}) denotes decile ranks. December Firms are firms with a December 31 year-end. Small (Large) Firms include those firms with market capitalization smaller (larger) than the first quintile of NYSE size distribution. Growth (Value) Firms include firms whose book-to-market is smaller (larger) than the median of NYSE book-to-market distribution. *, **, *** denote two-tail significance at the 10%, 5%, and 1% level, respectively.

Table 2.6: Characteristics of FFO/TL decile portfolios

	Portfolio FFO/TL ranking									
	lowest	2	3	4	5	6	7	8	9	highest
FFO/TL	−1.83	–0.14	0.04	0.11	0.15	0.20	0.26	0.35	0.51	1.23
CFO/ATA	–0.31	–0.03	0.05	0.08	0.10	0.12	0.14	0.16	0.17	0.20
Beta	1.46	1.28	1.13	1.07	1.05	1.05	1.05	1.10	1.11	1.20
ME	3.11	3.29	3.92	4.53	4.93	5.18	5.37	5.19	5.05	4.82
BE/ME	0.66	0.90	0.77	0.96	0.88	0.80	0.72	0.69	0.65	0.56
Ret1yr	7.20	12.27	13.90	16.50	17.24	16.66	17.33	18.30	16.40	14.78

Note to Table 2.6:

This table presents mean values of characteristics for ten portfolios formed annually at the end of June of year t from 1978 to 2007 by assigning firms to deciles based on the magnitude of FFO/TL_t. The sample consists of 74,364 firm-years representing 8,594 firms. *FFO/TL* is funds from operations divided by total liabilities and *CFO/ATA* is cash flows from operations divided by average total assets. Cash flows from operations are operating income after depreciation minus accruals. Accruals are defined as the change in non-cash current assets minus the change in current liabilities excluding the current portion of short-term debt and taxes payable minus depreciation expense. $Beta_t$ is estimated using the Capital Asset Pricing Model (CAPM) for each firm using monthly returns for at least 24 of the 60 months preceding July of year t. ME_t is log of a firm's market equity in millions at the end of June of year t. BE/ME_t is book common equity for the fiscal year ending in calendar year $t-1$, divided by market equity at the end of December $t-1$. $Ret1yr_t$ is the buy-and-hold returns (in percent) over one-year period beginning July of year t.

Table 2.7: Association of returns with CFO/ATA
Dependent variable = Ret1yr

	All Firms (N=74,364)	December Firms (N=38,691)	Small Firms (N=49,648)	Large Firms (N=24,716)	Growth Firms (N=43,662)	Value Firms (N=30,702)
Intercept	15.06***	14.96***	15.54***	14.37***	12.01***	19.35***
	(3.55)	(3.75)	(3.19)	(4.35)	(2.86)	(4.30)
CFO/ATA^{D}	11.04***	11.31***	11.39***	7.33***	10.77***	7.10**
	(4.27)	(3.62)	(3.43)	(4.06)	(3.48)	(2.75)
Beta^{D}	−0.04	0.28	1.36	−2.86	−0.96	0.98
	(−0.01)	(0.07)	(0.31)	(−0.61)	(−0.22)	(0.21)
ME^{D}	−3.29	−3.26	−2.96	0.75	−0.80	−6.72
	(−0.78)	(−0.79)	(−0.82)	(0.34)	(−0.20)	(−1.48)
BE/ME^{D}	14.39***	14.77***	15.43***	8.21**	12.57***	2.19
	(4.60)	(4.51)	(4.85)	(2.70)	(5.70)	(0.76)
Avg. Adj. R^2 (%)	4.50	5.04	3.74	6.57	5.31	2.34

Note to Table 2.7:

This table presents the results of Fama-MacBeth regressions using 74,364 firm-years (8,594 firms), spanning 30 years from 1978 to 2007. The dependent variable is regressed on scaled decile ranks, ranging from −0.5 to 0.5, of the independent variables each year. The regression coefficient reported here is the mean of the coefficient estimates over 30 sample years and *t*-statistic reported in parenthesis is the mean of the coefficient estimates divided by time-series standard error of the coefficient estimates. Variable definitions are in Table 2.6. Superscript D (^{D}) denotes decile ranks. December Firms are firms with a December 31 year-end. Small (Large) Firms include those firms with market capitalization smaller (larger) than the first quintile of NYSE size distribution. Growth (Value) Firms include firms whose book-to-market is smaller (larger) than the median of NYSE book-to-market distribution. *, **, *** denote two-tail significance at the 10%, 5%, and 1% level, respectively.

Table 2.8: Returns to zero-investment trading strategies using CFO/ATA

	Panel A: Equal-weighted portfolio			Panel B: Value-weighted portfolio		
Year	CFO/ATA Highest Decile Ret	CFO/ATA Lowest Decile Ret	Net Ret	CFO/ATA Highest Decile Ret	CFO/ATA Lowest Decile Ret	Net Ret
1978	19.77	15.54	4.24	17.68	17.81	−0.13
1979	21.17	11.99	9.18	9.32	17.66	−8.34
1980	41.45	47.58	−6.13	11.49	24.64	−13.15
1981	−16.90	−30.63	13.73	−10.11	−29.45	19.35
1982	86.50	98.15	−11.65	73.60	122.62	−49.02
1983	−11.85	−37.65	25.81	−8.70	−33.54	24.84
1984	20.91	−13.13	34.04	26.89	−11.83	38.72
1985	37.79	17.60	20.19	45.08	15.26	29.82
1986	16.34	−5.47	21.80	21.00	−15.49	36.49
1987	−7.33	−29.74	22.40	−10.70	−31.81	21.11
1988	13.40	−9.62	23.02	11.53	−7.36	18.89
1989	10.86	−5.36	16.21	20.44	−1.16	21.60
1990	9.48	6.85	2.63	22.20	−1.38	23.59
1991	14.05	37.96	−23.91	12.51	2.24	10.27
1992	22.68	21.93	0.75	1.38	−4.52	5.90
1993	1.47	−15.63	17.10	0.02	−21.30	21.32
1994	31.52	20.54	10.98	44.03	13.09	30.94
1995	29.96	47.50	−17.53	27.97	34.55	−6.58
1996	15.82	−32.74	48.56	38.90	−25.69	64.59
1997	15.63	−12.13	27.76	28.40	−18.18	46.58
1998	4.24	−5.38	9.62	25.78	2.32	23.47
1999	28.75	119.26	−90.51	42.24	104.21	−61.97
2000	12.45	−23.97	36.42	−36.85	−30.97	−5.88
2001	11.54	−40.15	51.70	−10.20	−48.65	38.45
2002	7.45	13.28	−5.83	6.49	22.72	−16.22
2003	36.81	30.03	6.78	12.51	26.66	−14.15
2004	14.40	−20.58	34.98	0.65	−20.57	21.22
2005	11.96	5.41	6.55	2.65	13.96	−11.31
2006	18.18	0.26	17.92	25.57	−3.74	29.31
2007	−5.88	−28.82	22.94	4.73	−20.74	25.46
Mean			10.99^{**}			12.17^{**}
(*t*)			(2.31)			(2.47)
Median			14.97^{***}			21.16^{***}
(signed rank)			(145.5)			(129.5)
No. of positive			$24/30^{***}$			$20/30^{*}$
(*Z*)			(3.29)			(1.83)

Note to Table 2.8:

This table presents returns to zero-investment trading strategies using cash flows from operations divided by average total assets (*CFO/ATA*) from 74,363 firm-years (8,594 firms), spanning 30 years from 1978 to 2007. At the end of June of each year *t*, firms are assigned into deciles according to their rankings of

CFO/ATA$_t$. *Ret* is the return on each portfolio (in percent) over one-year period beginning July of year *t*. *Mean* at the bottom of the *Net Ret* column is the average of the *Net Ret* figures. *, **, *** denote two-tail significance at the 10%, 5%, and 1% level, respectively.

Table 2.9: Associations of returns with FFO/TL before/after controlling for CFO/ATA
Dependent variable = Ret1yr

	All Firms (N=74,364)	December Firms (N=38,691)	Small Firms (N=49,648)	Large Firms (N=24,716)	Growth Firms (N=43,662)	Value Firms (N=30,702)
Panel A: Association of returns with FFO/TL before controlling for CFO/ATA						
Intercept	15.06***	14.96***	15.54***	14.37***	12.01***	19.35***
	(3.55)	(3.75)	(3.19)	(4.35)	(2.86)	(4.30)
FFO/TL$^{\mathbf{D}}$	8.07***	8.36***	9.64***	5.02**	8.30***	5.60**
	(4.38)	(3.84)	(3.36)	(2.62)	(4.32)	(2.16)
Beta$^{\mathbf{D}}$	−1.01	−0.60	0.86	−4.06	−2.08	0.85
	(−0.22)	(−0.14)	(0.19)	(−0.83)	(−0.43)	(0.18)
ME$^{\mathbf{D}}$	−1.86	−1.60	−2.91	1.63	0.60	−6.06
	(−0.42)	(−0.36)	(−0.83)	(0.71)	(0.14)	(−1.31)
BE/ME$^{\mathbf{D}}$	14.28***	14.68***	15.86***	7.09**	12.04***	2.63
	(4.59)	(4.45)	(4.87)	(2.46)	(5.36)	(0.92)
Avg. Adj. R^2 (%)	4.23	4.71	3.57	6.27	4.88	2.32
Panel B: Association of returns with FFO/TL after controlling for CFO/ATA						
Intercept	15.06***	14.96***	15.54***	14.37***	12.01***	19.35***
	(3.55)	(3.75)	(3.19)	(4.35)	(2.86)	(4.30)
FFO/TL$^{\mathbf{D}}$	3.51**	3.72**	5.13**	1.66	3.61**	3.50
	(2.25)	(2.27)	(2.42)	(0.69)	(2.06)	(1.45)
CFO/ATA$^{\mathbf{D}}$	9.14***	9.28***	8.69***	6.40**	8.51**	5.65**
	(3.44)	(2.93)	(2.94)	(2.75)	(2.42)	(2.35)
Beta$^{\mathbf{D}}$	0.03	0.48	1.61	−3.04	−1.06	1.22
	(0.01)	(0.12)	(0.38)	(−0.66)	(−0.24)	(0.27)
ME$^{\mathbf{D}}$	−3.72	−3.69	−3.64	0.79	−1.00	−6.87
	(−0.90)	(−0.92)	(−1.07)	(0.35)	(−0.25)	(−1.55)
BE/ME$^{\mathbf{D}}$	14.33***	14.80***	15.34***	7.99**	12.21***	2.74
	(4.62)	(4.51)	(4.97)	(2.58)	(5.38)	(0.96)
Avg. Adj. R^2 (%)	4.62	5.17	3.92	6.83	5.37	2.62

Note to Table 2.9:

This table presents the results of Fama-MacBeth regressions using 74,364 firm-years (8,594 firms), spanning 30 years from 1978 to 2007. The dependent variable is regressed on scaled decile ranks, ranging from −0.5 to 0.5, of the independent variables, each year. The regression coefficient reported here is the mean of the coefficient estimates over 30 sample years and *t*-statistic reported in parenthesis is the mean of the coefficient estimates divided by time-series standard error of the coefficient estimates. Variable definitions are in Table 2.6. Superscript D (^{D}) denotes decile ranks. December Firms are firms with a December 31 year-end. Small (Large) Firms include those firms with market capitalization smaller (larger) than the first quintile of NYSE size distribution. Growth (Value) Firms include firms whose book-to-market is smaller (larger) than the median of NYSE book-to-market distribution. *, **, *** denote two-tail significance at the 10%, 5%, and 1% level, respectively.

Table 2.10: Associations of returns with O-Score before/after controlling for CFO/ATA
Dependent variable = Ret1yr

	All Firms (N=74,364)	December Firms (N=38,691)	Small Firms (N=49,648)	Large Firms (N=24,716)	Growth Firms (N=43,662)	Value Firms (N=30,702)
Panel A: Association of returns with O-Score before controlling for CFO/ATA						
Intercept	15.06***	14.96***	15.54***	14.36***	12.01***	19.35***
	(3.55)	(3.75)	(3.19)	(4.35)	(2.87)	(4.30)
O-Score^{D}	−6.11***	−5.80**	−7.89***	−2.29	−5.90***	−2.92
	(−3.37)	(−2.51)	(−3.82)	(−0.99)	(−2.85)	(−1.27)
Beta^{D}	−1.70	−1.38	−0.04	−4.52	−2.90	0.29
	(−0.35)	(−0.30)	(−0.01)	(−0.94)	(−0.59)	(0.06)
ME^{D}	−2.39	−1.81	−3.56	1.57	0.34	−6.30
	(−0.52)	(−0.40)	(−1.02)	(0.68)	(0.08)	(−1.40)
BE/ME^{D}	13.17***	13.94***	14.53***	5.91**	11.63***	1.66
	(3.95)	(3.98)	(4.13)	(2.06)	(4.73)	(0.61)
Avg. Adj. R^2 (%)	4.05	4.54	3.31	6.29	4.72	2.13
Panel B: Association of returns with O-Score after controlling for CFO/ATA						
Intercept	15.06***	14.96***	15.54***	14.36***	12.01***	19.35***
	(3.55)	(3.75)	(3.19)	(4.35)	(2.86)	(4.30)
O-Score^{D}	−2.57	−2.11	−4.49**	0.25	−1.54	−1.61
	(−1.16)	(−0.80)	(−2.10)	(0.09)	(−0.57)	(−0.67)
CFO/ATA^{D}	10.24***	10.75***	9.94***	7.56***	10.24**	6.72**
	(3.42)	(3.06)	(2.77)	(3.45)	(2.73)	(2.50)
Beta^{D}	−0.06	0.46	1.43	−2.91	−1.10	1.03
	(−0.01)	(0.11)	(0.33)	(−0.66)	(−0.26)	(0.23)
ME^{D}	−4.28	−4.09	−4.36	1.02	−1.28	−7.10
	(−1.05)	(−1.03)	(−1.31)	(0.45)	(−0.33)	(−1.69)
BE/ME^{D}	13.82***	14.58***	14.37***	8.02**	12.15***	2.21
	(4.07)	(4.12)	(4.23)	(2.49)	(4.87)	(0.78)
Avg. Adj. R^2 (%)	4.65	5.29	3.91	7.07	5.47	2.53

Note to Table 2.10:

This table presents the results of Fama-MacBeth regressions using 74,364 firm-years (8,594 firms), spanning 30 years from 1978 to 2007. The dependent variable is regressed on scaled decile ranks, ranging from −0.5 to 0.5, of the independent variables, each year. The regression coefficient reported here is the mean of the coefficient estimates over 30 sample years and *t*-statistic reported in parenthesis is the mean of the coefficient estimates divided by time-series standard error of the coefficient estimates. Variable definitions are in Table 2.6, except *O-Score* whose definition is in Table 2.1. Superscript D (^{D}) denotes decile ranks. December Firms are firms with a December 31 year-end. Small (Large) Firms include those firms with market capitalization smaller (larger) than the first quintile of NYSE size distribution. Growth (Value) Firms include firms whose book-to-market is smaller (larger) than the median of NYSE book-to-market distribution. *, **, *** denote two-tail significance at the 10%, 5%, and 1% level, respectively.

Table 2.11: Returns across FFO/TL and O-Score quintile portfolios after controlling for CFO/ATA

Panel A: Returns across FFO/TL quintile portfolios after controlling for CFO/ATA

	CFO/ATA quintile					
FFO/TL quintile	lowest	2	3	4	highest	CFO/ATA highest – lowest
lowest	6.92	15.03	11.83	13.56	18.24	11.32***
	[318]	[94]	[38]	[22]	[23]	(3.59)
2	9.72	15.30	16.55	18.28	14.41	4.69*
	[87]	[172]	[133]	[71]	[34]	(1.71)
3	9.52	16.68	16.90	18.77	17.81	8.29**
	[39]	[103]	[148]	[142]	[64]	(2.07)
4	21.21	18.74	16.06	17.85	18.44	–2.77
	[26]	[71]	[103]	[156]	[140]	(–0.42)
highest	7.63	11.57	15.41	17.16	16.72	9.09***
	[26]	[56]	[74]	[105]	[235]	(3.17)
FFO/TL highest – lowest	0.71	–3.46	3.58	3.61	–1.52	
	(0.19)	(–1.18)	(0.75)	(0.80)	(–0.28)	

Panel B: Returns across O-Score quintile portfolios after controlling for CFO/ATA

	CFO/ATA quintile					
O-Score quintile	lowest	2	3	4	highest	CFO/ATA highest – lowest
lowest	10.34	12.65	16.07	15.95	16.39	6.06**
	[28]	[69]	[85]	[115]	[198]	(2.15)
2	15.53	16.96	15.32	17.40	17.35	1.82
	[39]	[89]	[112]	[136]	[120]	(0.41)
3	9.59	15.54	16.76	19.25	19.38	9.79***
	[62]	[112]	[123]	[118]	[80]	(3.22)
4	13.65	18.75	18.49	20.11	16.33	2.69
	[110]	[132]	[115]	[83]	[56]	(0.65)
highest	4.63	11.96	10.98	16.82	15.59	10.97***
	[256]	[93]	[61]	[44]	[41]	(3.05)
O-Score lowest – highest	5.71**	0.68	5.10	–0.88	0.80	
	(2.23)	(0.22)	(1.05)	(–0.19)	(0.26)	

Note to Table 2.11:

This table presents average annual returns (in percent) for each of 25 portfolios. Portfolios are formed according to the independent rankings of FFO/TL_t ($O\text{-}Score_t$) and CFO/ATA_t at the end of June of each year t from 1978 to 2007 in Panel A (Panel B). The average number of firms over the sample years for each of 25 portfolios is reported in bracket. The number in parenthesis represents a time-series t-statistic. *,

**, *** denote two-tail significance at the 10%, 5%, and 1% level, respectively.

Table 2.12: Sensitivity tests using post-SFAS No. 95 observations
Dependent variable = Ret1yr

	All Firms (N=47,948)	December Firms (N=26,448)	Small Firms (N=32,138)	Large Firms (N=15,810)	Growth Firms (N=27,845)	Value Firms (N=20,103)
Panel A: Repeat of Table 2.7 using post-SFAS No. 95 observations						
CFO_{SCF}/ATA^{D}	10.84***	9.75**	11.12**	7.92***	9.89**	7.85**
	(3.26)	(2.52)	(2.48)	(3.58)	(2.56)	(2.36)
Panel B: Repeat of Panel B in Table 2.9 using post-SFAS No. 95 observations						
FFO/TL^{D}	2.33	2.19	3.23	2.54	4.31	1.38
	(1.15)	(0.89)	(0.96)	(0.75)	(1.63)	(0.49)
CFO_{SCF}/ATA^{D}	9.26**	8.17*	8.93**	6.44*	6.73	7.01**
	(2.77)	(1.95)	(2.37)	(2.08)	(1.49)	(2.64)
Panel C: Repeat of Panel B in Table 2.10 using post-SFAS No. 95 observations						
$O\text{-}Score^{D}$	−2.09	0.47	−3.14	−0.69	−1.94	−0.43
	(−0.64)	(0.12)	(−1.08)	(−0.18)	(−0.47)	(−0.16)
CFO_{SCF}/ATA^{D}	10.10**	10.13*	9.75*	8.00**	9.11*	7.66**
	(2.43)	(2.05)	(1.95)	(2.70)	(1.78)	(2.20)

Note to Table 2.12:

This table presents the replication of Tables 2.7, 2.9, and 2.10 using post-SFAS No. 95 observations. The sample consists of 47,948 firm-years (6,309 firms), spanning 19 years from 1989 to 2007. The dependent variable is regressed on scaled decile ranks, ranging from −0.5 to 0.5, of the independent variables, each year. The regression coefficient reported here is the mean of the coefficient estimates over 19 sample years and *t*-statistic reported in parenthesis is the mean of the coefficient estimates divided by time-series standard error of the coefficient estimates. For brevity, only the results for the variables of interest are provided in the table. CFO_{SCF} is defined as cash flows from operations minus the cash portion of discontinued operations and extraordinary items. Definitions of the remaining variables are in Table 2.6, except *O-Score* whose definition is in Table 2.1. Subscript SCF ($_{SCF}$) denotes the statement of cash flows. Superscript D (^{D}) denotes decile ranks. December Firms are firms with a December 31 year-end. Small (Large) Firms include those firms with market capitalization smaller (larger) than the first quintile of NYSE size distribution. Growth (Value) Firms include firms whose book-to-market is smaller (larger) than the median of NYSE book-to-market distribution. *, **, *** denote two-tail significance at the 10%, 5%, and 1%

level, respectively.

Table 2.13: Sensitivity tests using Sloan's (1996) original accrual measure
Dependent variable = Ret1yr

	All Firms (N=74,364)	December Firms (N=38,691)	Small Firms (N=49,648)	Large Firms (N=24,716)	Growth Firms (N=43,662)	Value Firms (N=30,702)
Panel A: Association of returns with O-Score after controlling for ACC/ATA						
O-Score$^{\mathbf{D}}$	−6.93***	−6.81***	−8.91***	−2.59	−6.47***	−3.86
	(−3.71)	(−3.08)	(−4.41)	(−1.08)	(−3.09)	(−1.61)
ACC/ATA$^{\mathbf{D}}$	−3.26*	−4.15**	−3.71*	−2.51	−1.70	−5.36**
	(−1.97)	(−2.44)	(−1.96)	(−1.67)	(−1.05)	(−2.15)
Panel B: Association of returns with ACC/ATA after controlling for CFO/ATA						
ACC/ATA$^{\mathbf{D}}$	3.43	2.13	4.06	1.73	4.24*	−0.80
	(1.28)	(0.71)	(1.28)	(0.75)	(1.88)	(−0.18)
CFO/ATA$^{\mathbf{D}}$	12.94***	12.57***	13.64***	8.13***	12.81***	6.86
	(3.66)	(3.05)	(3.16)	(3.06)	(3.57)	(1.46)
Panel C: Association of returns with O-Score after controlling for ACC/ATA and CFO/ATA						
O-Score$^{\mathbf{D}}$	−1.68	−1.67	−3.95	0.88	0.42	−2.53
	(−0.51)	(−0.51)	(−1.19)	(0.26)	(0.11)	(−0.78)
ACC/ATA$^{\mathbf{D}}$	2.70	1.44	2.42	1.91	4.35	−1.87
	(0.76)	(0.41)	(0.59)	(0.63)	(1.45)	(−0.36)
CFO/ATA$^{\mathbf{D}}$	12.04**	11.76**	11.54*	8.71**	13.05**	5.73
	(2.47)	(2.26)	(2.04)	(2.34)	(2.56)	(1.02)

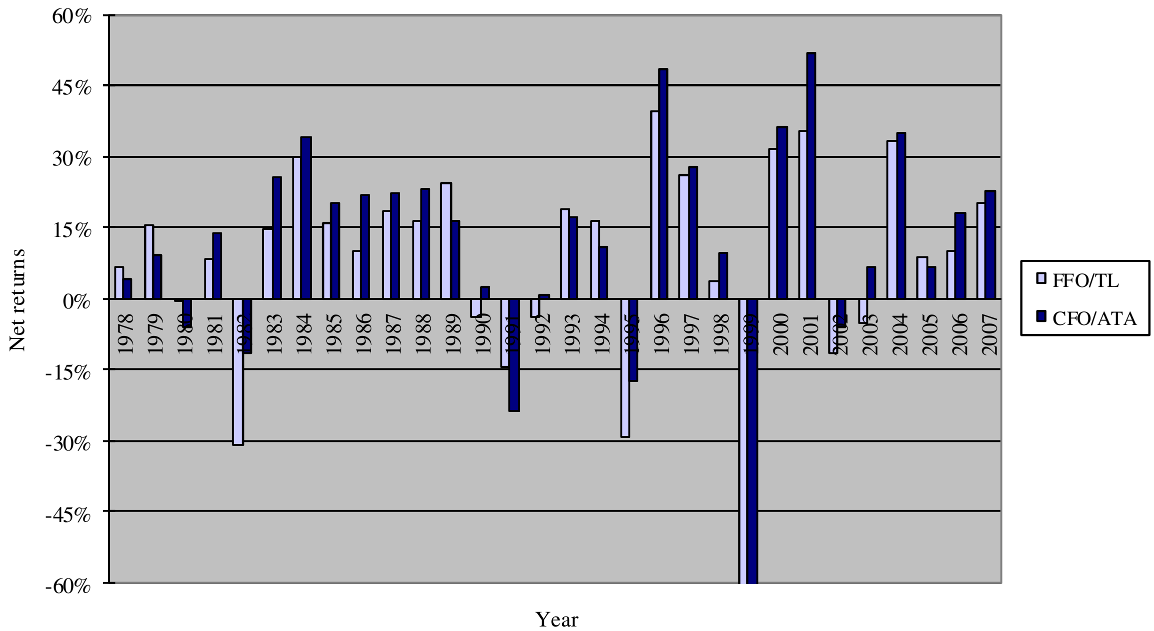
Note to Table 2.13:

This table presents the results of sensitivity tests using Sloan's (1996) original accrual measure. The results provided are based on Fama-MacBeth regressions using 74,364 firm-years (8,594 firms), spanning 30 years from 1978 to 2007. Beta, ME, and BE/ME are included in the regression but their results are not reported here for brevity. The dependent variable is regressed on scaled decile ranks, ranging from −0.5 to 0.5, of the independent variables, each year. The regression coefficient reported here is the mean of the coefficient estimates over 30 sample years and t-statistic reported in parenthesis is the mean of the coefficient estimates divided by time-series standard error of the coefficient estimates. Variable definitions are in Tables 1 and 6, except ACC/ATA, which represents accruals divided by average total assets, where accruals are defined as the change in non-cash current assets minus the change in current liabilities excluding the current portion of long-term debt and taxes payable minus depreciation expense. Superscript

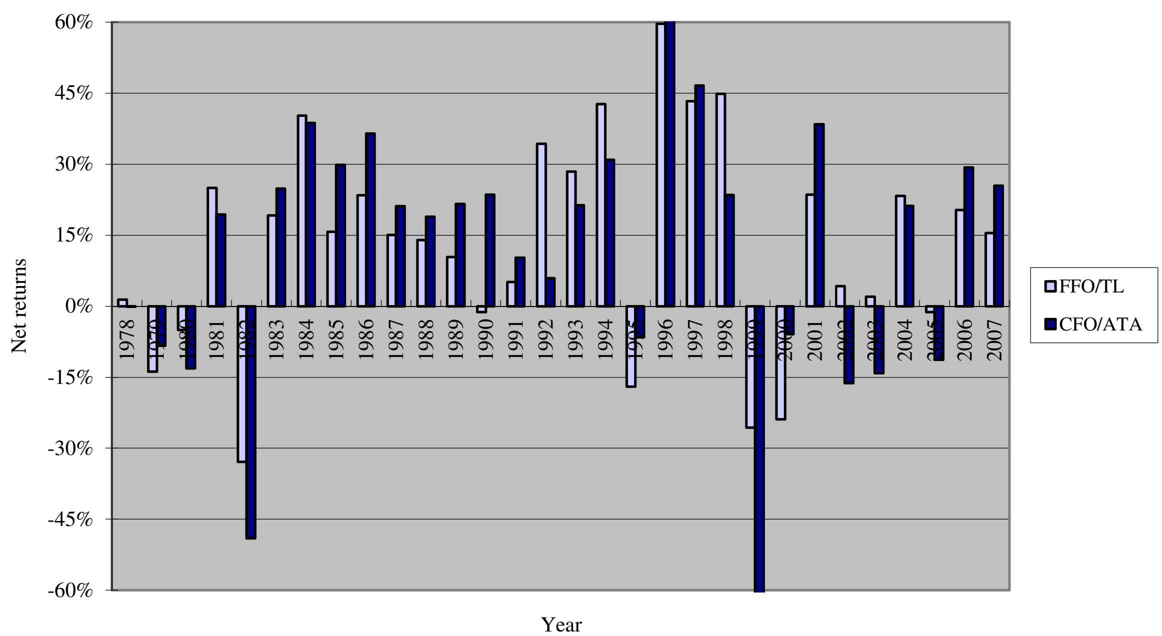
D (^{D}) denotes decile ranks. December Firms are firms with a December 31 year-end. Small (Large) Firms include those firms with market capitalization smaller (larger) than the first quintile of NYSE size distribution. Growth (Value) Firms include firms whose book-to-market is smaller (larger) than the median of NYSE book-to-market distribution. *, **, *** denote two-tail significance at the 10%, 5%, and 1% level, respectively.

Figure 2.1: Returns to zero-investment trading strategy using FFO/TL and CFO/ATA

Panel A: Equal-weighted portfolio



Panel B: Value-weighted portfolio



Chapter 3: Implementability of trading strategies based on accounting information: Piotroski (2000) revisited

3.1 Introduction

Chapter 2's investigation of the bankruptcy risk anomaly makes use of standard research design elements that are common to tests of return predictability. Generally, this "anomalies literature" explores whether firm characteristics such as book-to-market, bankruptcy risk measures, ratio of accruals to total assets, etc., can be used to identify mispriced stocks.

Although the literature is now fairly voluminous, there are still differences in opinion as to why detected mispricing persists. One explanation for detected mispricing is that excess returns represent compensation for risk (e.g., French, Schwert, and Stambaugh, 1987; Chan, 1998; Ball and Kothari, 1989; Kim and Kim, 2003; Khan, 2008). As noted in Fama (1970), this explanation is particularly difficult to prove or disprove because market efficiency must be tested jointly with a model for expected returns: in order to test whether returns are anomalous, one must know what expected returns should be in the absence of mispricing. Skeptics have also explored the degree to which excess returns documented by researchers are obtainable in practice due to institutional constraints such as short-sale constraints (e.g., Jones and Lamont, 2002; Boehme, Danielsen, and Sorescu, 2006) or transaction costs (e.g., Bhushan, 1994; Mashruwala, Rajgopal, and Shevlin, 2006; Ng, Rusticus, and Verdi, 2008). Finally, some critics point to research design choices or flaws that over- or understate the degree to which excess returns can be earned (e.g., Lo and MacKinlay, 1990; Kothari, Shanken, and Sloan, 1995; Kothari, Sabino, and Zack, 2005; Kraft, Leone, and Wasley, 2006).

This study contributes to this last research stream. It investigates the effects of two potentially problematic research design choices that are often made in accounting-based studies of anomalies. Both of these research design choices are relevant to the class of studies that employ zero-investment hedge portfolios to measure the economic significance of anomalous returns and do not apply to the alternative methodology that relates, via regression, future returns to current information. The two implementation issues are: 1) the use of equal- versus value-weights in forming the portfolios and 2) the selection of return accumulation periods based on firm specific fiscal year-ends, rather than establishing a common investment date for all firms in the hedge portfolio.

I will argue that equal weighting of stocks within the hedge portfolios can be called into question. This first issue, coupled with transaction costs, implies that the returns earned on equally weighted stocks within an assumed portfolio generated for research purposes are not likely to mimic real-world returns on the same set of stocks. The second issue revolves around the fact that if a money manager were to sort firms on the variable of interest such as book-to-market she cannot determine the total number of firms in each side of the hedge portfolio until the accounting information of the entire portfolio becomes available. Therefore, she is unable to decide the applicable weight for an individual firm within the portfolio. For this reason, I label the implementability issue related to the return accumulation period as the problem of unknown portfolio weights.

This study explores these two issues by re-examining the results in Piotroski (2000). That study investigates whether a simple, accounting-based fundamental analysis strategy can enhance the returns to so-called value investing (i.e., investing in high book-to-market firms) documented by Fama and French (1992) and Lakonishok, Shleifer, and Vishny (1994), among others. Piotroski's study is motivated by the observation that while the value investing strategy, on average, yields positive abnormal returns, the success of that strategy relies on the strong performance of a small fraction of firms and thus, value investors must tolerate poor performance from a large fraction of the firms in their portfolios. Given the diverse outcomes realized within high book-to-market firms, Piotroski asks whether a simple, financial statement-based heuristic, when applied to a subset of firms with high book-to-market ratios, can discriminate between the firms that will eventually provide high returns and those that will be poor performers.

To identify the firms that will eventually be strong value firms, Piotroski (2000) establishes a trading heuristic that is based on nine fundamental signals to measure the financial condition of value firms. The nine signals are based on widely used measures of profitability, financial leverage/liquidity, and operating efficiency.[20] He classifies each firm's realized signal as either "good" or "bad" depending on the signal's implication for future performance. He assigns one (zero) to an indicator variable if the signal's realization is good (bad) and then defines the aggregate signal measure, F-Score, as the sum of the nine binary signals. With a subset of firms with high book-to-market ratios, he finds that firms with high F-Scores earn higher average subsequent returns than those with low F-Scores during his sample period of 1976 to 1999.

It is worth noting that Piotroski (2000) is not the only study that is subject to the two issues of trading strategy implementability raised in this study. However, there are additional reasons that make revisiting Piotroski a worthy exercise. First, if the purpose of a study is to enhance the value investing strategy, then the methodology used in the study should be consistent with prior studies that document excess returns from value investing. Piotroski does not follow this general principle. Specifically, all three finance studies cited in Piotroski on value investing (i.e., Rosenberg, Reid, and Lanstein, 1984; Fama and French, 1992; and Lakonishok, Shleifer, and Vishny, 1994) start accumulating returns of individual firms at a common date (i.e., July 1 of each year). In contrast, Piotroski accumulates returns starting four months after each firm's fiscal year-end. Thus, the question arises, would he have obtained similar results even if he had followed the approach used in these finance studies upon which he relied? Second, Pi-

[20] Piotroski identifies nine fundamental signals based on his own reading of the professional and academic articles, not based on statistical techniques such as factor analysis.

otroski's research question leads him to employ a subset of firms that are likely to be particularly vulnerable to the portfolio weighting issue raised in this study.

To investigate the impact of the unknown portfolio weights problem, I replicate the analysis in Piotroski (2000), which I refer to as Analysis 1, and then adjust the return accumulation period such that returns of individual firms are accumulated with a common starting date, which I refer to as Analysis 2. To isolate the impact of this problem on Piotroski's results, equal weighting is applied to both analyses. The comparison of returns to zero-investment trading strategies reveals that the trading strategy in Analysis 1 outperforms that in Analysis 2 in 22 out of 32 years from 1976 to 2007. The difference in mean and median are significantly different from zero. Taken together, my results suggest that the return accumulation methodology typically used in accounting studies is likely to result in inflated excess returns relative to the one used in the finance literature.

To investigate the effect of portfolio weighting on the evidence of market inefficiency, I apply value weighting to Analysis 2, where the weight assigned to a firm is determined based on its market capitalization at the beginning of return accumulation date. In contrast to Piotroski (2000), which finds a mean annual return of 9.70 percent over 21 years from 1976 to 1996, I find that the F-Score-based zero-investment strategy only yields a mean annual return of 2.49 percent, insignificantly different from zero, over 32 years from 1976 to 2007. I also find that firms with high F-Scores only outperform those with low F-Scores in 20 out of 32 years and the accompanying Z-statistic shows that the proportion of positive returns is not likely to be greater than 50 percent. Taken as a whole, my results highlight that when value-weighted returns are examined, the evidence of stock market inefficiency shrinks significantly and becomes statistically unreliable. Thus, they remind researchers of the importance of a value-weighting scheme as a robustness check at a minimum, if not the focus of the main analysis.

3.2 Background and motivation

3.2.1 Equal- versus value-weighting

Most accounting studies that challenge market efficiency, including Piotroski (2000), exclusively rely on equal-weighting when forming portfolios. Equal-weighting by itself does not pose any econometric problems. However, if the objective is to show market inefficiency by generating profits from a trading strategy, then the adopted methodology should reflect the economic circumstances under which trades are executed in practice.

Academic researchers often use the NYSE-Amex-NASDAQ universe to maximize sample size and statistical power. However, the majority of firms in this universe are small firms, calling into question the real-world potential of implementing an equal-weighting scheme. According to Fama and French (2008), microcaps, which they define as stocks with market capitalization below the 20th NYSE percentile, are on average about 3 percent of the NYSE-Amex-NASDAQ universe in value, but they account for about 60 percent of the total number of stocks. Although researchers implicitly assume that there would have been a large enough supply of shares at a particular price at a particular point in time, this assumption is more likely to be violated if the shares under consideration are from small firms because these shares tend to be illiquid. Furthermore, documented profits from trading strategies are reported gross of transaction costs. These costs are important in determining the performance of the strategies and can substantially reduce documented returns to an investment strategy (Keim and Madhavan, 1997). While an equal-weighting scheme implicitly assumes equal transaction costs across firms, this is not the case in the real world. For example, Lesmond, Ogden, and Trzcinka (1999) estimate that average round-trip transaction costs from 1963 to 1990 are 1.2 percent and 10.3 percent for the largest and smallest NYSE-Amex decile firms, respectively.

Prior research in finance has proposed value-weighting as a superior alternative to equal-weighting. Nonetheless, most accounting studies that challenge market efficiency have exclusively relied on equal-weighting when forming portfolios. In response to this, Lewellen (2010), in his review of recent empirical work on market inefficiency, proposes that "empirical studies should focus on value-weighted returns in addition to (or instead of) equal-weighted returns (p. 462)."

3.2.2 The problem of unknown portfolio weights

Accounting studies that raise doubts about the informational efficiency of stock markets tend to document positive returns on zero-investment portfolios constructed on the basis of publicly available information. These studies typically accumulate returns for individual firms starting three or four months after their fiscal year-ends. The gap between fiscal year-end and return accumulation period starting date ensures that (1) the financial accounting information of a firm becomes publicly available before its returns are accumulated and (2) accounting information used is not stale. Finance studies, meanwhile, accumulate returns with a common starting date each year regardless of the fiscal year-end of individual firms. For example, Lakonishok, Shleifer, and Vishny (1994) in their study of the value anomaly accumulate returns starting at the end of April each year for all observations with fiscal years ending in the prior calendar year and thus, the gap between fiscal year-end and the common starting date is four months for firms with December fiscal year-ends, but this gap could be as much as 15 months for firms with January fiscal year-ends.

Regression analyses that relate future returns to current accounting information are also commonly used to identify anomalies because one can extract the marginal effects of the variable of interest using multiple regression slopes. As long as researchers limit their analyses to regression, the approach used in accounting studies with respect to return accumulation does not pose any problem. However, researchers oftent extend their analyses into portfolio analyses (1) to provide a simple picture of how returns vary across the spectrum of the variables of interest and (2) to facilitate gauging the economic significance of identified anomalies. The typical approach in portfolio analyses is to form decile portfolios by sorting stocks on the variable of interest. Though returns for individual deciles are typically shown, it is common to focus on the hedge portfolio return obtained from long-short positions in the extreme deciles. This is where the problem of unknown portfolio weights arises.

Hedging long-short positions requires the same amount of investment in each position. To achieve this, one has to know how many and which firms belong to each position, which, in turn, determines the amount of investment in individual firms in each position. If a money manager were to sort firms on the variable of interest such as book-to-market, she cannot determine the number and identity of firms in each side of the hedge portfolio until the accounting information of the entire portfolio becomes available. Therefore, she cannot decide an applicable weight for an individual firm within the portfolio. I illustrate the problem of unknown portfolio weights in the context of Piotroski (2000) to enhance understanding of this problem.

Suppose there are 100 firms in the market of interest and that 40 of these have June fiscal year-ends and the remainder have December year-ends. For the sake of simplicity, assume each firm can only have an F-Score of zero or nine. A money manager will go long on firms with F-Scores of nine and short on firms with F-Scores of zero. She will take a total of $3,000 in each position and assign an equal weight to individual firms in each position. At the end of April of year *t+1*, accounting information for all 100 firms becomes available and thus, the fund manager can determine the number of individual firms within each F-Score group. This enables her to decide the amount of investment in the individual firms as shown in Figure 3.1. At the end of October of year *t*, by contrast, she has information on F-Scores of only the first 40 firms (those with June fiscal year-ends). This makes it impossible to decide the amount of investment in the individual firms because the total number of individual firms in each side of the hedge portfolio is not yet available. Therefore, the analysis suffers from the problem of unknown portfolio weights if a researcher assumes knowledge of the number of individual observations within a hedge portfolio before the accounting information of the entire portfolio becomes publicly available.

3.2.3 Survey of prior accounting studies related to accounting anomalies

This section provides a focused survey of recent archival empirical literature related to accounting anomalies because this study contributes to that literature to the extent that the implementability issues discussed are widespread in the accounting literature.[21] My survey focuses on research studies published in *The Accounting Review*, *Journal of Accounting and Economics*, and *Journal of Accounting Research* within a 10-year period starting with the year 2000. Although my survey is limited only to studies published in three journals, the methodologies employed in these studies are often adopted by subsequent studies in other journals. Therefore, I believe my survey captures a broad picture of the methods used in recent studies on anomalies.

I was able to identify 14 studies, including Piotroski (2000), that are related to accounting

---

[21] The focus of my survey is to identify the implementability issues raised in this study. Refer to Richardson, Tuna, and Wysocki (2010) for a more general survey of recent studies on anomalies.

anomalies. Of these 14 studies, three studies limit their analysis to regressions and thus, they are not subject to the implementability issues discussed in this study as noted above.[22] Table 3.1 provides information about the remaining 11, including (1) portfolio formation methodology and (2) whether the particular study is subject to the problem of unknown portfolio weights.

The second column of Table 3.1 confirms that unlike finance studies, in which value-weighted hedge portfolio returns are often shown along with equal-weighted returns, a significant proportion of accounting studies rely exclusively on equal-weighting. Among the articles identified, Hirshleifer, Hou, Teoh, and Zhang (2004) is the only study that provides value-weighted portfolio returns as well as equal-weighted returns. Equal-weighting by itself does not pose any econometric problems, but equal-weighted returns do not accurately capture the total wealth effects experienced by investors. Moreover, money managers are unlikely to be able to replicate an anomaly derived from equal-weighting.

Whether the use of value-weighted returns would change the tenor of results concerning the remaining 10 studies remains an open question, but Fama (1998) provides a helpful insight as to the robustness of anomalies derived from equal-weighted returns. After re-examining anomalies documented in the finance literature, he concludes that

---

[22] These three studies are Ali, Hwang, and Trombley (2003); Fairfield, Whisenant, and Yohn (2003); and Pincus, Rajgopal, and Venkatachalam (2007).

> "Whenever value-weighted returns are examined, apparent anomalies shrink a lot and typically become statistically unreliable. At a minimum, this suggests that anomalies are largely limited to small stocks. But a reasonable alternative explanation is that small stocks are just a sure source of bad-model problems (p. 304)."

Thus, Fama's findings not only challenge the validity of 10 accounting studies that only rely on equal-weighting but also call for a robustness check using value-weighted returns.

As shown in the third column of Table 3.1, only three of 11 studies are free of the problem of unknown portfolio weights. These studies avoid this problem by limiting their samples to firms with December fiscal year-end (Beneish and Vargus, 2002; Mashruwala, Rajgopal, and Shevlin, 2006) or forming portfolios in every month (Hirshleifer, Hou, Teoh, and Zhang, 2004). Of the eight studies that are subject to the problem of unknown portfolio weights, three address this problem in their sensitivity analysis. For example, Xie (2001) states that

> "My sample includes both December fiscal year-end and non-December fiscal year-end firms. Accounting information for a given fiscal year will become available to the market at different points in calendar time for firms with different fiscal year-end months. Thus, one cannot directly implement the hedge-portfolio strategies reported in Table 3 (p. 367)."

After acknowledging the problem, these studies repeat their analysis by forming portfolios in every month (Xie) or by limiting their sample to firms with December fiscal year-end (Penman and Zhang, 2002; Desai, Rajpogal, Venkatachalam, 2004).

Thus, it is clear that many studies published in the three surveyed journals are subject to the issues of implementability of trading strategies raised in this study. Given the influence of these journals on the literature, additional studies published in other journals can be expected to have the same implementability issues as well. Thus, the implications of my findings are not limited to the case of Piotroski (2000).

3.3 Description of the empirical tests

To investigate the impact of the unknown portfolio weights problem, I replicate the analysis in Piotroski (2000), which I refer to as Analysis 1, and then change the return accumulation period such that returns of individual firms are accumulated with a common starting date, which I refer to as Analysis 2. To isolate the impact of this problem on Piotroski's results, equal weighting is applied to both analyses. In the subsequent analysis, I apply value weighting to Analysis 2 to investigate the effect of portfolio weighting on the evidence of market inefficiency. Before I lay out the details of Analysis 1 and Analysis 2, I discuss how Piotroski constructs his F-Score using fundamental analysis.

3.3.1 F-Score

Piotroski (2000) introduces F-Score to differentiate between financially strong and weak firms among value stocks. Specifically, he chooses nine fundamental signals to measure three areas of a firm's financial condition: profitability, financial leverage/liquidity, and operating efficiency. A firm's F-Score is the sum of nine binary signals and provides a measure of the strength of the firm's financial position. Piotroski classifies each firm's signal realization as either "good" or "bad" depending on the signal's implication for future performance. An indicator variable for the signal is equal to one (zero) if the signal's realization is good (bad). Thus, a high (low) F-Score represents a firm with mostly (very few) good signals.

Piotroski (2000) employs four measures of profitability because these measures provide information about the firm's ability to generate funds internally: ROA, CFO, ΔROA, and ACCRUAL. He defines these variables as follows:

ROA_t	=	net income before extraordinary items in year *t* scaled by total assets at the beginning of year *t*;
CFO_t	=	cash flows from operations in year *t* scaled by total assets at the beginning of year *t*;
ΔROA_t	=	ROA in year *t* less ROA in year $t-1$;
$ACCRUAL_t$	=	$ROA_t - CFO_t$.

For ROA, CFO, and ΔROA, he assigns one to F_ROA, F_CFO, and F_ΔROA, respectively, if each signal is positive and zero otherwise. He assigns one to F_ACCRUAL if ACCRUAL is negative and zero otherwise because earnings driven by positive accrual adjustments have been found to be a bad signal about future profitability (Sloan, 1996).

To measure changes in capital structure and the firm's ability to meet future debt obligations, Piotroski (2000) uses these signals: ΔLIQUID, ΔLEVER, and EQ_OFFER, defined as follows:

$\Delta LIQUID_t$ = change in the ratio of total current assets to total current liabilities between the end of year $t-1$ and year t;

$\Delta LEVER_t$ = change in the ratio of total long-term debt to total assets between the end of year $t-1$ and year t;

EQ_OFFER_t = an indicator variable equal to one if a firm did not issue common equity in the year t, zero otherwise.

Piotroski assigns one to F_ΔLIQUID if ΔLIQUID is positive and zero otherwise because an improvement in liquidity is expected to be a good signal about the firm's ability to service short-term debt obligation. He assigns one to F_ΔLEVER if ΔLEVER is negative and zero otherwise. As noted in Piotroski, an increase in leverage itself is not necessarily a negative signal with respect to future performance, but he reasons that because value firms tend to be financially distressed to some extent they are signaling their inability to generate sufficient internal funds by issuing debt (Myers and Majluf, 1984; Miller and Rock, 1985). Using similar reasoning, Piotroski views issuing common equity as a bad signal and thus assigns one to EQ_OFFER when a firm does not issue common equity.[23]

---

[23] To identify equity offerings, Piotroski depends on COMPUSTAT data108 (SSTK in the fundamental file), sale of common and preferred stock, which includes 1) sale of common shares, 2) sale of preferred shares, 3) conversion of preferred shares and/or debt into common shares, and 4) exercise of stock options and/or warrants. Thus, a positive value in the above data item does not necessarily mean that a company has issued common equity. Nonetheless, I follow Piotroski because 1) this study is replicating Piotroski and 2) the impact of the measurement error is expected to be marginal because the main analysis in this study combines firms into two groups.

Piotroski (2000) uses two signals to measure changes in the efficiency of the firm's operations: ΔMARGIN and ΔTURN, defined as follows:

ΔMARGIN_t = change in the gross margin ratio between year $t-1$ and year t;

ΔTURN_t = change in the asset turnover ratio between year $t-1$ and year t.

Piotroski assigns one to F_ΔMARGIN and F_ΔTURN, respectively, if each signal is positive and zero otherwise because an improvement in either ratio suggests an improvement in operational efficiency.

Having defined these nine signals, F-Score is then simply the sum of these signals:

$$\begin{aligned}\text{F-Score} = {} & \text{F\_ROA} + \text{F\_CFO} + \text{F\_\_}\Delta\text{ROA} + \text{F\_ACCRUAL} \\ & + \text{F\_}\Delta\text{LIQUID} + \text{F\_}\Delta\text{LEVER} + \text{EQ\_OFFER} \\ & + \text{F\_}\Delta\text{MARGIN} + \text{F\_}\Delta\text{TURN}.\end{aligned}$$

3.3.2 Details on Analysis 1 and Analysis 2

Analysis 1 intends to closely replicate Piotroski (2000) and Analysis 2 intends to vary the return accumulation period such that returns of individual firms are accumulated with a common starting date.

I perform the following steps for Analysis 1:

Step 1: I calculate book-to-market quintile cutoffs for each calendar year using all firms with sufficient data within the merged COMPUSTAT/CRSP annual fundamental file. To calculate the market value of equity, I use a firm's fiscal year-end share price and the number of shares outstanding at the fiscal year-end.

Step 2: I calculate F-Score for all observations in the merged COMPUSTAT/CRSP annual fundamental file.

Step 3: I merge cutoffs from Step 1 and the dataset in Step 2 such that cutoffs in year $t-1$ are assigned to the dataset in year t. Then, I retain only firm-years in the highest book-to-market quintile. If F-Score for a particular firm-year is missing due to the lack of required financial statement data, then the observation is dropped from the sample.

Step 4: I calculate one-year market-adjusted returns for all available observations using the CRSP monthly stock file, where market-adjusted return is defined as the buy-and-hold return less the value-weighted market return.

Step 5: Using the cross-reference file linking COMPUSTAT's GVKEY to the historical perm number on CRSP (LPERMNO), I merge the dataset derived from Step 3 and the dataset from Step 4 such that firm-years have one-year market-adjusted returns starting at the end of the sixth month after the fiscal year-end.

Although Analysis 1 closely follows Piotroski (2000), there are a few differences. First, my approach combines firm-years with fiscal year-ends falling on the same calendar year as a unit of yearly analysis, whereas Piotroski defines years according to FYEAR in COMPUSTAT, which ends on May 31. This choice was made in order to allow for comparison with Analysis 2. Second, Analysis 1 accumulates returns for individual observations starting six months after their fiscal year-end, rather than four months used in Piotroski. I intentionally switch to a six-month gap based on the findings of Alford, Jones, and Zmijewski (1994) that 2 percent of the 10-K reports are not filed by the end of the fifth month after the fiscal year-end. Another reason for switching to a six-month gap is that Piotroski does not limit his sample to U.S. firms, so his sample includes a number of ADRs, which have six months after the fiscal year-end to file Form 20-F.

Note that the final sample does not include firms with negative book-to-market ratios because I use only firms in the highest book-to-market quintile. Nonetheless, the decision of whether to include negative book-to-market firms in Step 1 is not trivial in that it affects the quintile cut-offs used to determine the composition of the final sample. Piotroski (2000) is silent about the inclusion (or exclusion) of negative book-to-market firms with respect to the calculation of cutoffs, but the replication done in this study suggests that firms with negative book-to-market ratios are included in his study. Hence, I include firms with negative book-to-market ratios in Step 1. Also note that I use the delisting return in the delisting month in Step 4. If a delisting is due to liquidation or a forced delisting by either the exchange or the Securities and Exchange Commission (SEC) (CRSP delisting codes 400 to 699) and the delisting return is missing, the delisting return is set to negative 100 percent.[24] If a firm delists within one year after return accumulation begins, then the one-year market-adjusted return is defined as the buy-and-hold return through the delisting month less the value-weighted market return over the corresponding time period.

---

[24] By contrast, Piotroski assumes that a missing delisting return is zero. While I do not agree with Piotroski's decision here, I find that setting the delisting return to zero does not change the tenor of my results quantitatively or qualitatively.

Analysis 2 requires the following steps (differences from Analysis 1 are underlined).

Step 1: I calculate book-to-market quintile cutoffs for each calendar year using all firms with sufficient data within the merged COMPUSTAT/CRSP annual fundamental file. To calculate the market value of equity, I use a firm's fiscal year-end price and the number of shares outstanding at the fiscal year-end (the same as Step 1 in Analysis 1).

Step 2: I calculate F-Score for all observations in the merged COMPUSTAT/CRSP annual fundamental file (the same as Step 2 in Analysis 1).

Step 3: I merge cutoffs from Step 1 and the dataset in Step 2 such that cutoffs in year *t* are assigned to the dataset in year *t*. Then, I retain only firm-years in the highest book-to-market quintile. If F-Score for a particular firm-year is missing due to the lack of required financial statement data, then the observation is dropped from the sample.

Step 4: I calculate one-year market-adjusted returns for all available observations using the CRSP monthly stock file, where market-adjusted return is defined as the buy-and-hold return less the value-weighted market return (the same as Step 4 in Analysis 1).

Step 5: Using the cross-reference file linking COMPUSTAT's GVKEY to the historical perm number on CRSP (LPERMNO), I merge the dataset derived from Step 3 and the dataset from Step 4 such that firm-years have one-year market-adjusted returns starting at the end of June of the following year.

Note that Analysis 2 employs the current year book-to-market distribution in Step 3 because necessary information to determine the current year distribution is available at the beginning of the applicable return accumulation period. Because this study retains only firms in the highest book-to-market quintile, the use of different cutoffs in Step 3 across Analysis 1 and Analysis 2 leads to the creation of different samples across the two analyses, which will be discussed in more detail in the next section.

3.4 Sample

The empirical tests in this study are conducted using all firms in the intersection of the 2010 editions of the merged COMPUSTAT/CRSP annual fundamental file and the CRSP monthly stock file. To be included in the sample, (1) the firm-year must have the accounting information necessary to calculate its F-Score, (2) the firm-year's book-to-market falls in the highest book-to-market quintile, and (3) the share price and number of shares outstanding are available from CRSP at the beginning of the applicable return accumulation period. The third requirement implies that included firms survived until the beginning of the applicable return accumulation period. For firms that fail after the return accumulation period begins, the delisting returns are used as explained above.

Consistent with Piotroski (2000), the sample period in this study begins in 1976. Due to the need to accumulate returns up to 30 months following the end of a fiscal year, the sample period ends in 2007. For example, firm-years with fiscal year-ends falling in January to December 2007 are grouped together and their returns are accumulated from July 2008 to June 2009. Taking into account the data restrictions discussed in Section 3.3, the final sample for Analysis 1 (Analysis 2) is 22,520 (23,118) firm-years representing 6,429 (6,297) firms. The intersection of the two samples is 19,889 firm-years representing 5,898 firms.

Discrepancies between the two samples in the number of observations arise from the requirements of Step 3 and Step 5. With regard to Step 3, Analysis 1 uses the previous year's book-to-market distribution, whereas Analysis 2 uses the current year's distribution. Thus, the increase or decrease in the highest quintile cutoff from year $t-1$ to year t results in some observations being included in the final sample of one but not the other analysis.[25] With respect to Step 5, there are two sources of differences in the final sample composition. First, Analysis 1 requires firms to survive until the end of six months after the fiscal year-end, whereas Analysis 2 requires firms to survive longer, with the exception of firms with December fiscal year-ends. Thus, if a firm delists between the beginning of the return accumulation period of Analysis 1 and that of Analysis 2, then the firm is only included in the sample for Analysis 1. Second, if a firm's shares are suspended from trading at the beginning of the return accumulation period of Analysis 1 but trading is resumed on or before that of Analysis 2, the firm is only included in the sample of Analysis 2.

To summarize the differences in the two samples, I decompose the final samples into those that are common across Analysis 1 and Analysis 2 and those that are unique to one or the

---

[25] For example, assume the highest book-to-market quintile cutoff in year 2001 was 1.0 and the corresponding figure in year 2002 was 1.5. Then, firm-years with a book-to-market that falls between 1.0 and 1.5 in year 2002 will only be included in Analysis 1. On the other hand, if the cutoff declined from 1.5 in year 2001 to 1.0 in year 2002, then firm-years with book-to-market between 1.0 and 1.5 in year 2002 will only be included in Analysis 2.

other analysis. I further decompose each subsample based on F-Score. As can be seen in Table 3.2, all observations are heavily concentrated in the intermediate F-Scores. This holds true for the observations that are common to both samples as well as those that are unique to one sample. This result is expected because F-Score is a sum of random variables with finite means and variances, so the distribution of F-Score should approach normal as the number of observations increases (Central Limit Theorem).

Because the main empirical tests are performed on a yearly basis, differences in sample composition by year are also investigated. The results are shown in Table 3.3. The highest book-to-market quintile cutoff used in this study declines over my sample period from 1.83 in 1976 to 0.83 in 2007. This is consistent with Givoly and Hayn (2000), who find that the average book-to-market has decreased since the 1970s. Although there is an overall decreasing trend in book-to-market, the trend is not entirely monotonic year over year: out of the 32 years examined in this study, the quintile cutoff decreased 22 times from year *t-1* to year *t*, but it also increased 10 times. The differences in sample composition reported in Table 3.3 are quite dramatic in some years because of a unique set of firm-years belonging only to the sample for Analysis 1 or that for Analysis 2. For example, in 1976 the number of firm-years in the sample for Analysis 2 is almost twice as large as that for Analysis 1.

3.5 Empirical results

3.5.1 The problem of unknown portfolio weights

The main analyses start with replicating Table 3 of Piotroski (2000, p. 17): I calculate mean and median one-year market-adjusted returns of each F-Score group for the 32-year period spanning from 1976 to 2007 and report the results in Table 3.4.

Consistent with Piotroski (2000), mean and median one-year market-adjusted returns are monotonically increasing in F-Score, including the exception of firms with F-Scores of zero. In addition, firms with F-Scores of eight or nine deliver economically and statistically higher returns than those with F-Scores of zero or one. In Table 3 of Piotroski, where the sample period spans 21 years from 1976 to 1996, mean and median difference in market-adjusted returns between firms with F-Scores of 8 or 9 and those with F-Scores of zero or one are 23.0 and 20.0 percent, respectively. The results in Table 3.4 are comparable to those of Piotroski: mean and median differences are 20.59 and 24.56 percent, respectively. The results for Analysis 2, where the problem of unknown portfolio weights in Analysis 1 is eliminated, provide insight into whether unknown portfolio weights would impact conclusions drawn from Analysis 1. Comparing the results of Analysis 1 to those of Analysis 2, it is clear that returns of firms with F-Scores of five or greater decline with regard to both mean and median, but returns of firms with F-Scores of four or less provide mixed results. Nonetheless, the difference between firms with F-Scores of eight or nine and those with F-Scores of zero or one is still economically and statistically significantly different from zero whether I focus on Analysis 1 or Analysis 2.

In his main analysis, Piotroski (2000) presents the gap between returns of firms with F-Scores of eight or nine and those with F-Scores of zero or one as evidence of stock market inefficiency. However, the gap does not represent an excess return one could obtain over the period of 32 years from 1976 to 2007 by going long on firms with F-Scores of eight or nine and going short on firms with F-Scores of zero or one. The excess return and the corresponding statistical significance in Table 3.4 are calculated as if observations coexist in one year. However, that is not the case because the data spans 32 years. Thus, it is impossible to determine how much should be invested in each stock to form a zero-investment portfolio. Furthermore, the feasibility of the trading strategy that employs firms with F-Scores of eight or nine and those with F-Scores of zero or one is to some extent doubtful: there is no guarantee that firms with extreme scores will exist in a particular year.

Following Piotroski's (2000) yearly analysis and to address these two concerns, Tables 3.5 and 3.6 present results by year. I classify all firms in a particular year into two groups: one with F-Scores of five through nine and another with F-Scores of zero through four. As in Piotroski, I call the former Strong and the latter Weak for brevity. For each year from 1976 to 2007, I form equally-weighted portfolios for Weak and Strong under Analysis 1, as well as Analysis 2, and report the associated returns in Table 3.5. I also report the performance of All, which represents all firms in both Strong and Weak. By construction, performance of All is the weighted average of the performances of Strong and Weak. The performance of Weak is not significantly different from that of the market index across Analysis 1 and Analysis 2. Strong clearly outperforms the market index as evidenced by the tests of mean, median, and proportion in each analysis.

Table 3.6 formally investigates the impact of unknown portfolio weights. It reports and compares the profits of hedge portfolios derived under Analysis 1 and Analysis 2, using equally weighted returns in the hedge portfolios. Given the presence of short-sale constraints (Jones and Lamont, 2002; Boehme, Danielsen, and Sorescu, 2006), it is also important to compare the performance of Strong and that of All: the former represents the outcome of value investing refined with F-Score and the latter represents the outcome of value investing as originally conceived. Thus, I also report Strong minus All, in addition to Strong minus Weak. Under Analysis 1, the mean (median) annual return on Strong minus Weak is 8.33 percent with *t*-statistic of 6.22 (8.00 percent with significance at the 1 percent level). The profitability of the strategy is consistent: Strong outperforms Weak in 27 out of 32 years. Although the mean and median annual returns of

Strong minus All are smaller, they show similar statistical significance.[26]

As explained above, the returns reported under Analysis 1 rely on accounting information that is not available at the time of portfolio formation. Analysis 2 effectively removes this concern but with a trade-off of incorporating stale information. Implementing Analysis 2 reveals that the profitability of hedge portfolios under Analysis 1 (mean of 8.33 and median of 3.24 percent) is more than 40 percent higher relative to that under Analysis 2 (mean of 5.80 and median of 2.26 percent). The comparison of the trading strategy in the last two columns shows that the trading strategy in Analysis 1 produces higher returns than that in Analysis 2 in 22 out of 32 years and the differences in mean and median are significantly different from zero. Thus, Table 3.6 demonstrates that the relationship between Piotroski's (2000) fundamental signals and subsequent returns is attenuated if the portfolios are formed only after knowledge of the portfolio weights becomes available using a common date to accumulate returns, as is done in many studies in the finance literature.

---

[26] Note that the sign of Strong minus All should be the same as that of Strong minus Weak, and the magnitude of Strong minus All is equal to one minus the proportion of Strong to total firms multiplied by the magnitude of Strong minus Weak by construction. The numerical relationship between Strong minus All and Strong minus Weak is as follows:

$$\text{Strong} - \text{All} = \text{Strong} - [\alpha*\text{Strong} + (1 - \alpha)*\text{Weak}] = (1 - \alpha)*(\text{Strong} - \text{Weak}),$$

where α is the proportion of Strong to total firms and $0 < \alpha < 1$.

3.5.2 Equal- versus value-weighting

The results provided in Table 3.6 show that the unknown portfolio weights problem attenuates, but does not eliminate, the positive relationship between F-Score and subsequent returns. However, the profit generated using the trading strategy is potentially unobtainable due to the issue of equal-weighting discussed earlier. As noted above, equal-weighting does not reflect the economic circumstances in which trades are executed. Thus, I revisit Analysis 2 with a value-weighting scheme.

Table 3.7 reports one-year market-adjusted returns to value-weighted portfolios for each year from 1976 to 2007. This table is identical to Panel B of Table 3.5 with the exception of portfolio weighting. The weight assigned to a firm is determined based on its market capitalization at the beginning of the return accumulation date (i.e., July 1 each year). For brevity, I limit my discussion to the Strong portfolios because Table 3.5 shows that most of the excess returns derive from this side of the hedge portfolios. The mean annual return on Strong is 5.21 percent with *t*-statistic of 2.09, while the median annual return is 4.85 percent with significance at the 10 percent level. These figures are about half of the corresponding figures in Panel B of Table 3.5. Moreover, as shown in the bottom rows of Table 3.7, Strong outperforms the market index in only 18 of 32 years and the accompanying *Z* statistic shows that the proportion of positive market-adjusted returns is not significantly greater than 50 percent.

Table 3.8 provides a formal test of excess hedge portfolio returns. As shown in the bottom rows, the Strong minus Weak strategy only yields a mean annual return of 2.49 percent, insignificantly different from zero. While the median annual return is slightly higher than the mean return, it is also not significantly different from zero. Furthermore, Strong only outperforms Weak in 20 of 32 years and the accompanying Z statistic shows that the proportion of positive returns is not significantly greater than 50 percent. The results in Table 3.8 thus demonstrate that firms with high F-Scores do not outperform those with low F-Scores when value-weighting is employed and the problem of unknown portfolio weights is corrected.

Table 3.8 also compares returns to hedge portfolios under Analysis 1 with equal-weighting to those under Analysis 2 with value-weighting. Mean returns to hedge portfolios under the former are greater than those under the latter by 5.84 percent with t-statistic of 2.75. This difference represents the combined impact of the unknown portfolio weights problem and the change in portfolio weighting, demonstrating that these research design choices can have substantial consequences when exploring evidence of stock market inefficiency. Because Analysis 1 with equal-weighting is typical of those often employed in the accounting literature, the results in Table 3.8 suggest that evidence of profitable trading strategies and market inefficiency is likely to be overstated in a substantial number of studies.

3.5.3 Real-world implementation of F-Score-based trading strategy

Analysis 2, the approach used in finance studies, provides a solution to the unknown portfolio weights problem but with a trade-off of using stale information in portfolio formation. Thus, this approach is unlikely to be the approach used by money managers. Instead, they would implement a trading strategy multiple times, possibly every month, to better use current information. For example, by increasing the frequency of portfolio formation under Analysis 2 from once a year to twice a year, one can reduce the maximum gap between fiscal year-end and the beginning of the return accumulation period by six months (from 17 months to 11 months). In the extreme, one can entirely avoid incorporating stale information by forming portfolios every month. This section revisits the F-Score-based trading strategy by implementing the strategy multiple times a year.[27]

I start by implementing an F-Score-based trading strategy twice a year and report returns to hedge portfolios in Panel A of Table 3.9. To better reflect the economic circumstances in which trades are executed, I form value-weighted portfolios. Each year firms with fiscal year-end from January 1 to June 30 (July 1 to December 31) are grouped together and their returns are accumulated starting January 1 (July 1) of the following year. The mean annual return on the hedge portfolios is 2.70 percent with *t*-statistic of 1.30 and the median annual return is 4.54 per-

[27] Another way to avoid incorporating stale information is to limit a sample to firms with December fiscal year-ends. When I repeat the analysis using only firms with December fiscal year-ends, profits to hedge portfolios are not materially different from those obtained under value-weighting in Table 3.8: mean (median) returns are 4.30 percent with *t*-statistic of 1.63 (5.22 percent with significance at the 10 percent level) and the proportion of positive returns is 21 out of 32 with *Z*-statistic of 1.77.

cent with significance at the 10 percent level. Compared to the results obtained under value-weighting in Table 3.8, results in Panel A suggest that implementing the strategy multiple times a year does not improve the profitability of hedge portfolios.

I further increase the frequency of portfolio formation to four times a year and six times a year. Technically speaking, I could form portfolios every month, but I decide not to because of the small number of firms in the Weak group. Implementing the strategy four times a year produces results that are similar to implementing the strategy twice a year. When the frequency is increased to six times a year, there is a marginal improvement in the profitability of the hedge portfolios. The mean annual return on the hedge portfolios is 3.38 percent with *t*-statistic of 1.71 and the median annual return is 2.36 percent with significance at the 5 percent level. The profitability of the trading strategy is marginally consistent: the proportion of positive returns is 109 of 192 iterations with *Z*-statistic of 1.88. However, the reported profits are not large enough to offset related transaction costs. Thus, the information in Table 3.9 suggests that Piotroski's (2000) trading strategy is unlikely to generate a reliable profit even if it is implemented by a money manager.

3.5.4 Replication of Xie (2001)

Many prior accounting studies using trading strategies are subject to the issues of implementability raised in this study. I replicate and extend Piotroski (2000) because his research design employs a subset of firms that are likely to be more vulnerable to the issue of implementability.

While this makes the test setting advantageous, this is a double-edged sword: I may not find similar results when I repeat the analyses with other subsets of firms or when I apply the analyses to other anomalies. To answer concerns about the true impact of the issues I have identified on studies other than Piotroski, I investigate the implementability issues in the context of Xie (2001).

Xie (2001) is an extension of Sloan's (1996) accrual anomaly. Xie estimates abnormal accruals with a cross-sectional version of the Jones (1991) model (DeFond and Jiambalvo, 1994; Subramanyam, 1996). Then he shows that the overpricing of total accruals documented by Sloan is due largely to abnormal accruals. I replicate Xie because this is the paper with the highest average citations per year among the 54 anomalies-related papers published in accounting journals in the last 10 years (Richardson, Tuna, and Wysocki, 2010).

In replicating Xie (2001), I limit my sample to firms with fiscal year-ends falling between 1971 and 1992 so that my sample period coincides with that of Xie. I closely follow Xie's research design, but I make a few modifications. The gap between fiscal year-end and the beginning of the return accumulation period is three months in Xie, whereas I allow a six-month gap, as in my main analysis. Second, Xie uses COMPUSTAT definition of FYEAR for the year identification (year *t* includes firm-years whose fiscal year-end falls between June 1, year *t* and May 31, year *t+1*), whereas I use the calendar year to be consistent with my main analysis. Third, Xie excludes NASDAQ observations prior to 1982 due to the lack of size-adjusted returns in the

1995 CRSP Indices file, whereas I include these observations because the 2010 CRSP Indices file supplies size-adjusted returns for these firms. Xie's final sample consists of 56,692 firm-years (representing 7,506 firms), whereas my final sample consists of 68,549 firm-years (representing 8,225 firms).

Each year the hedge portfolio is formed by taking a long position in the lowest decile portfolio and a short position in the highest decile portfolio based on abnormal accruals. Annual returns to the hedge portfolios are reported in Table 3.10, which is directly comparable to Panel A of Table 3 of Xie (2001, p. 366). The mean return over the 22 year period from 1971 to 1992 in Table 3.10 is 7.50 percent with *t*-statistic of 7.16 and the corresponding figure in Xie is 11.0 percent with *t*-statistic of 8.43. The lowest decile outperforms the highest decile in 21 of 22 years in Table 3.10, whereas Xie reports that the lowest decile outperforms the highest decile in all 22 years. Overall, results provided in Table 3.10 are similar to those reported in Xie.

As in my main analysis, I accumulate returns of individual firms starting July 1 of each year in order to eliminate the problem of unknown portfolio weights, and then I form equal- and value-weighted portfolios. When equal-weighted portfolios are formed, the mean and median return and the proportion of positive returns becomes smaller to some extent but they are still statistically significant at the 1 percent level. This is consistent with what I find in the context of Piotroski (2001) and also consistent with what Xie (2001) reports in his sensitivity test, in which he forms equal-weighted portfolios in every month.

The previously discussed analyses in Tables 3.7 to 3.9 show that value weighting diminishes excess returns. In contrast, hedge portfolios formed based on abnormal accruals with value weighting still generate statistically significant profits, although the statistical significance is moderated. The mean annual return on the hedge portfolios is 5.56 percent with *t*-statistic of 2.16 and the median annual return is 6.59 percent with significance at the 10 percent level. The most striking difference when my results are compared to Xie's (2001) is the proportion of positive returns. When value-weighted portfolios are formed (after eliminating the problem of unknown portfolio weights), the profitability of the trading strategy becomes much less consistent: the lowest decile outperforms the highest decile in only 15 of 22 years with *Z*-statistic of 1.71. Although the combined impact of the unknown portfolio weights problem and different portfolio weighting on Xie is not as dramatic as that on Piotroski (2000), Table 3.10, taken as a whole, demonstrates that the implementation issues raised in this study are not limited to Piotroski.

3.6 Conclusions

This study investigates two aspects relating to the implementability of trading strategies used in studies on accounting-related anomalies. First, the return accumulation approach used in accounting studies cannot be replicated in a practical context because the number and identity of individual observations within a portfolio are assigned within a research context before the accounting information of all firms in the portfolio would actually be available in real time. Second, the fact that the majority of the firms in the NYSE-Amex-NASDAQ universe are small impedes the implementation of an equal-weighting scheme. As discussed, equal weighting has been the approach used in most accounting studies.

I examine these two issues by replicating and extending the analysis in Piotroski (2000). I find that the relationship between Piotroski's fundamental signals and subsequent returns is partly driven by the choice of return accumulation periods and the use of equally weighted returns. When the research design controls for both problems, the relationship between F-Scores and subsequent returns disappears. Because the methods used in Piotroski are typical of those often employed in the accounting literature, my study suggests that evidence of profitable trading strategies and market inefficiency in the literature is likely to be overstated.

This study contributes to a stream of research that proposes research design choices or flaws as an explanation of detected mispricing (e.g., Lo and MacKinlay, 1990; Kothari, Shanken, and Sloan, 1995; Kothari, Sabino, and Zack, 2005; Kraft, Leone, and Wasley, 2006). I demonstrate that two research design choices commonly made in accounting studies can bias the evidence in favor of finding stock market inefficiency. Of the two research design choices I examine, it appears that the decision to equal-weight rather than value-weight stocks within a portfolio can be the most influential.

Table 3.1: Implementability of trading strategies in prior accounting studies

Authors (Year published) [Publication outlet]	Portfolio formation methodology	Subject to problem of unknown portfolio weights?*
Piotroski (2000) [Journal of Accounting Research]	Equal-weighting only	Yes
Xie (2001) [The Accounting Review]	Equal-weighting only	Yes#
Beneish and Vargus (2002) [The Accounting Review]	Equal-weighting only	No
Penman and Zhang (2002) [The Accounting Review]	Equal-weighting only	Yes#
Desai, Rajgopal, and Venkatachalam (2004) [The Accounting Review]	Equal-weighting only	Yes#
Hirshleifer, Hou, Teoh, and Zhang (2004) [Journal of Accounting and Economics]	Equal-weighting and value-weighting	No
Richardson, Sloan, Soliman, and Tuna (2005) [Journal of Accounting and Economics]	Equal-weighting only	Yes
Mashruwala, Rajgopal, and Shevlin (2006) [Journal of Accounting and Economics]	Equal-weighting only	No
Penman, Richardson, and Tuna (2007) [Journal of Accounting Research]	Equal-weighting only	Yes
Zhang (2007) [The Accounting Review]	Equal-weighting only	Yes
Dechow, Richardson, and Sloan (2008) [Journal of Accounting Research]	Equal-weighting only	Yes

Note to Table 3.1:

This table compiles the results of a survey of recent archival empirical literature related to accounting anomalies. The survey focuses on research studies published in *The Accounting Review*, *Journal of Accounting and Economics*, and *Journal of Accounting Research* within a 10-year period starting the year 2000. # denotes a case where a sensitivity test in a particular study addresses problem of unknown portfolio weights.

* I classify a particular study as being subject to the problem of unknown portfolio weights when (1) the study accumulates returns for individual firms starting three or four months after their fiscal year-ends *and* (2) it does not limit its sample to firms with December fiscal year-end.

Table 3.2: The sample composition across Analysis 1 and Analysis 2 by F-Score

F-Score	Unique to Analysis 1: Due to increase in cutoff from year t-1 to year t	Unique to Analysis 1: Due to subsequent delisting	Common to Analysis 1 and Analysis 2	Unique to Analysis 2: Due to decrease in cutoff from year t-1 to year t	Unique to Analysis 2: Due to suspension of trading
0	11	1	56	7	0
1	64	16	553	66	0
2	150	46	1,593	211	4
3	308	85	2,768	351	6
4	412	112	3,546	528	4
5	401	102	3,816	586	3
6	369	74	3,283	630	8
7	230	52	2,486	485	2
8	141	22	1,428	256	5
9	29	6	360	76	1
Total	2,115	516	19,889	3,196	33

Note to Table 3.2:

This table reports the number of observations by F-Score in the final samples of Analysis 1 and Analysis 2. Spanning 32 years from 1976 to 2007, the final sample of Analysis 1 (Analysis 2) is 22,520 (23,118) firm-years representing 6,429 (6,297) firms. In Analysis 1, following Piotroski (2000), returns of individual observations are accumulated starting six months after the fiscal year-end, whereas returns under Analysis 2 are accumulated with a common starting date (i.e., July 1 each year), which is consistent with the approach used in finance studies. Analysis 1 (Analysis 2) employs firm-years in the highest quintile of the previous (current) year's book-to-market distribution. Thus, the increase (decrease) in the highest quintile cutoff from year $t-1$ to year t results in unique observations in the final sample of Analysis 1 (Analysis 2). There are two additional sources of differences in the sample composition. First, Analysis 1 requires firms to survive until the end of six months after the fiscal year-end, whereas Analysis 2 requires firms to survive longer, with the exception of firms with December fiscal year-ends. Thus, if a firm delists between the beginning of the return accumulation period of Analysis 1 and that of Analysis 2, then the firm is only included in the sample for Analysis 1. Second, if a firm's shares are suspended from trading at the beginning of the return accumulation period of Analysis 1 but trading is resumed on or before that of Analysis 2, the firm is only included in the sample of Analysis 2. F-Score represents a summary measure of a firm's financial condition introduced by Piotroski and it is the sum of the following nine binary signals: (1) F_ROA is equal to one if ROA is positive, and zero otherwise, where ROA is net income before extraordinary items scaled by total assets at the beginning of the year, (2) F_CFO is equal to one if CFO is positive, and zero otherwise, where CFO is cash flows from operations scaled by total assets at the beginning of the year, (3) F_ΔROA is equal to one if ΔROA is positive, and zero otherwise, where ΔROA is current year ROA minus prior year ROA, (4) F_ACCRUAL is equal to one if ACCRUAL is negative, and zero otherwise, where ACCRUAL is ROA minus CFO, (5) F_ΔLIQUID is equal to one if ΔLIQUID is positive, and zero otherwise, where ΔLIQUID is change in the ratio of total current assets to total current liabilities from prior year to current year, (6) F_ΔLEVER is equal to one if ΔLEVER is negative, and zero otherwise, where ΔLEVER is change in the ratio of total long-term debt to total assets from prior year to current year, (7) EQ_OFFER is equal to one if a firm does not issue common equity, and zero otherwise, (8) F_ΔMARGIN is equal to one if ΔMARGIN is positive, and zero otherwise, where ΔMARGIN is change in the gross margin ratio from prior year to current year, and (9) F_ΔTURN is equal to one if ΔTURN is positive, and zero otherwise, where ΔTURN is change in the asset turnover ratio from prior year to current year.

Table 3.3: The sample composition across Analysis 1 and Analysis 2 by year

Year	Unique to Analysis 1: Due to increase in cutoff from year t-1 to year t	Unique to Analysis 1: Due to subsequent delisting	Common to Analysis 1 and Analysis 2	Unique to Analysis 2: Due to decrease in cutoff from year t-1 to year t	Unique to Analysis 2: Due to suspension of trading
1976	0	4	329	298	0
1977	0	10	564	50	1
1978	0	3	544	63	3
1979	0	8	550	63	1
1980	0	12	592	34	2
1981	0	12	595	57	1
1982	0	18	607	42	4
1983	0	4	242	426	3
1984	118	23	727	0	1
1985	0	18	585	155	3
1986	0	15	653	76	1
1987	174	9	687	0	0
1988	0	19	660	9	0
1989	0	21	640	14	0
1990	438	10	645	0	2
1991	0	23	507	175	0
1992	0	11	514	249	3
1993	0	12	564	109	0
1994	112	12	677	0	0
1995	0	17	684	102	0
1996	0	22	775	155	0
1997	0	28	801	177	0
1998	450	29	904	0	0
1999	180	23	853	0	0
2000	276	32	826	0	1
2001	0	23	702	161	1
2002	59	28	895	0	1
2003	0	17	376	441	0
2004	0	11	469	262	0
2005	53	18	701	0	2
2006	0	16	526	78	2
2007	255	8	495	0	1
Total	2,115	516	19,889	3,196	33

Note to Table 3.3:

This table reports the number of observations by year in the final samples of Analysis 1 and Analysis 2. Spanning 32 years from 1976 to 2007, the final sample of Analysis 1 (Analysis 2) is 22,520 (23,118) firm-years representing 6,429 (6,297) firms. In Analysis 1, following Piotroski (2000), returns of individual observations are accumulated starting six months after the fiscal year-end, whereas returns under

Analysis 2 are accumulated with a common starting date (i.e., July 1 each year). Analysis 1 (Analysis 2) employs firm-years in the highest quintile of the previous (current) year's book-to-market distribution. Thus, the increase (decrease) in the highest quintile cutoff from year $t-1$ to year t results in unique observations in the final sample of Analysis 1 (Analysis 2). There are two additional sources of differences in the sample composition. First, Analysis 1 requires firms to survive until the end of six months after the fiscal year-end, whereas Analysis 2 requires firms to survive longer, with the exception of firms with December fiscal year-ends. Thus, if a firm delists between the beginning of the return accumulation period of Analysis 1 and that of Analysis 2, then the firm is only included in the sample for Analysis 1. Second, if a firm's shares are suspended from trading at the beginning of the return accumulation period of Analysis 1 but trading is resumed on or before that of Analysis 2, the firm is only included in the sample of Analysis 2.

Table 3.4: One-year market-adjusted returns by F-Score

	Analysis 1				Analysis 2			
F-Score	mean	median	n	n%	mean	median	n	n%
0	4.64%	−24.67%	68	0.3%	13.85%	−29.02%	63	0.3%
1	−4.10	−20.50	633	2.8%	−7.46	−20.02	619	2.7%
2	4.27	−12.82	1,789	7.9%	0.54	−13.70	1,808	7.8%
3	9.36	−8.81	3,161	14.0%	6.60	−10.22	3,125	13.5%
4	8.46	−6.97	4,070	18.1%	8.11	−5.96	4,078	17.6%
5	11.06	−2.95	4,319	19.2%	8.88	−2.84	4,405	19.1%
6	13.85	−0.38	3,726	16.5%	11.29	−1.74	3,921	17.0%
7	14.76	0.87	2,768	12.3%	12.18	0.30	2,973	12.9%
8	15.83	1.91	1,591	7.1%	13.78	1.91	1,689	7.3%
9	23.41	6.70	395	1.8%	13.74	3.98	437	1.9%
All	10.84	−4.36	22,520	100%	8.64	−4.36	23,118	100%
F-Score 1&2 [A]	−3.26	−21.40	701	3.1%	−5.49	−21.59	682	3.0%
F-Score 8&9 [B]	17.34	3.17	1,986	8.8%	13.78	2.29	2,126	9.2%
[B] − [A]	20.59***	24.56***			19.27***	23.89***		

Note to Table 3.4:

This table presents one-year market-adjusted returns (in percent) over the period of 1976-2007 by F-Score, where the market-adjusted return is the buy-and-hold return less the value-weighted market return. F-Score represents a summary measure of a firm's financial condition introduced by Piotroski (2000). See Table 3.2 for more details on F-Score. In Analysis 1, following Piotroski, returns of individual observations are accumulated starting six months after the fiscal year-end, whereas returns under Analysis 2 are accumulated with common starting date (i.e., July 1 each year). Differences in mean and median are tested with two-sample *t*-test and signed rank test, respectively. *, **, *** denote two-tail significance at the 10%, 5%, and 1% level, respectively.

Table 3.5: One-year market-adjusted returns to F-Score-based equal-weighted portfolios

Panel A: Returns to Weak, Strong, and All portfolios under Analysis 1

	Strong portfolios		Weak portfolios		All portfolios	
Year	n	one year market-adjusted return	n	one year market-adjusted return	n	one year market-adjusted return
1976	261	33.32%	72	28.74%	333	32.33%
1977	404	21.61	170	15.53	574	19.81
1978	376	–0.50	171	–11.03	547	–3.80
1979	372	22.64	186	6.18	558	17.15
1980	420	13.41	184	0.15	604	9.37
1981	426	35.34	181	24.67	607	32.16
1982	382	30.30	243	20.99	625	26.68
1983	174	–1.70	72	–26.38	246	–8.92
1984	563	–4.65	305	–19.06	868	–9.72
1985	323	6.52	280	–11.82	603	–2.00
1986	351	10.68	317	2.69	668	6.89
1987	524	6.35	346	–4.37	870	2.09
1988	392	–9.57	287	–15.91	679	–12.25
1989	359	–11.00	302	–7.94	661	–9.60
1990	614	20.32	479	16.05	1093	18.45
1991	306	26.43	224	18.42	530	23.04
1992	339	32.75	186	18.67	525	27.76
1993	333	14.22	243	11.08	576	12.90
1994	453	2.22	348	2.93	801	2.53
1995	396	–2.65	305	–17.01	701	–8.90
1996	441	–0.33	356	–12.67	797	–5.84
1997	451	–12.35	378	–15.18	829	–13.64
1998	665	18.64	718	28.51	1383	23.76
1999	535	22.51	521	8.69	1056	15.69
2000	604	35.50	530	28.07	1134	32.03
2001	344	12.29	381	15.51	725	13.98
2002	479	59.49	503	61.33	982	60.43
2003	208	23.18	185	–1.18	393	11.71
2004	274	15.45	206	6.51	480	11.61
2005	405	10.05	367	3.48	772	6.93
2006	282	–9.06	260	–16.23	542	–12.50
2007	343	–4.89	415	–9.40	758	–7.36
Mean		13.02***		4.69		9.77***
(*t*)		(4.37)		(1.41)		(3.19)
Median		12.85***		3.21		10.49***
(signed rank)		(189)		(64)		(150)
No. of positive		22/32**		19/32		21/32*
(*Z*)		(2.12)		(1.06)		(1.77)

Table 3.5: One-year market-adjusted returns to F-Score-based equal-weighted portfolios (cont.)

Panel B: Returns to Weak, Strong, and All portfolios under Analysis 2

	Strong portfolios		Weak portfolios		All portfolios	
Year	n	one year market-adjusted return	n	one year market-adjusted return	n	one year market-adjusted return
1976	502	38.79%	125	35.57%	627	38.15%
1977	427	9.98	188	–0.71	615	6.71
1978	416	–5.47	194	–10.04	610	–6.92
1979	410	30.06	204	23.26	614	27.80
1980	433	11.68	195	0.67	628	8.26
1981	462	45.89	191	44.65	653	45.53
1982	397	9.32	256	1.94	653	6.43
1983	480	–2.71	191	–24.13	671	–8.81
1984	465	–3.50	263	–13.54	728	–7.13
1985	400	9.27	343	–5.04	743	2.66
1986	390	9.44	340	0.26	730	5.16
1987	409	2.94	278	–4.74	687	–0.17
1988	389	–8.78	280	–19.80	669	–13.39
1989	356	–2.64	298	–7.51	654	–4.86
1990	349	17.04	298	17.23	647	17.13
1991	406	31.61	276	12.48	682	23.87
1992	494	18.68	272	16.51	766	17.91
1993	389	6.33	284	5.14	673	5.83
1994	383	2.24	294	9.40	677	5.35
1995	440	–9.62	346	–25.77	786	–16.73
1996	522	–5.14	408	–6.88	930	–5.90
1997	540	–19.39	438	–15.95	978	–17.85
1998	444	12.17	460	26.72	904	19.58
1999	425	33.00	428	14.52	853	23.73
2000	432	38.65	395	27.09	827	33.13
2001	411	11.92	453	12.31	864	12.12
2002	439	51.68	457	45.41	896	48.48
2003	468	8.70	349	6.10	817	7.59
2004	443	11.90	288	9.96	731	11.14
2005	366	7.97	337	5.21	703	6.64
2006	322	–12.91	284	–19.55	606	–16.02
2007	216	–8.92	280	–6.28	496	–7.43
Mean		10.63***		4.83		8.37**
(*t*)		(3.38)		(1.47)		(2.69)
Median		9.30***		3.54		6.53**
(signed rank)		(155)		(64)		(118)
No. of positive		22/32**		19/32		21/32*
(*Z*)		(2.12)		(1.06)		(1.77)

Note to Table 3.5:

This table presents one-year market-adjusted returns (in percent) to equal-weighted portfolios, where portfolios are formed based on F-Score. F-Score represents a summary measure of a firm's financial condition introduced by Piotroski (2000). See Table 3.2 for more details on F-Score. Strong (Weak) represents firms with F-Scores of 5 or greater (0 through 4) and All represents all firms. In Analysis 1, following Piotroski, returns of individual observations are accumulated starting six months after the fiscal year-end, whereas returns under Analysis 2 are accumulated with a common starting date (i.e., July 1 each year). Differences in mean and median are tested with two-sample *t*-test and signed rank test, respectively. *, **, *** denote two-tail significance at the 10%, 5%, and 1% level, respectively.

Table 3.6: The impact of the unknown portfolio weights problem on the profitability of hedge portfolios

Year	Piotroski (2000) replicated — Analysis 1 under equal-weighting: Strong – Weak [A]	Strong – All [B]	Unknown portfolio weights problem eliminated — Analysis 2 under equal-weighting: Strong – Weak [A*]	Strong – All [B*]	Impact of unknown portfolio weights problem: [A] – [A*]	[B] – [B*]
1976	4.58%	0.99%	3.22%	0.64%	1.36%	0.35%
1977	6.08	1.80	10.68	3.27	–4.61	–1.47
1978	10.53	3.29	4.57	1.45	5.96	1.84
1979	16.46	5.49	6.80	2.26	9.67	3.23
1980	13.26	4.04	11.01	3.42	2.25	0.62
1981	10.67	3.18	1.24	0.36	9.44	2.82
1982	9.32	3.62	7.38	2.89	1.94	0.73
1983	24.68	7.22	21.41	6.10	3.27	1.13
1984	14.41	5.06	10.04	3.63	4.37	1.44
1985	18.34	8.52	14.32	6.61	4.02	1.91
1986	7.99	3.79	9.18	4.28	–1.19	–0.48
1987	10.73	4.27	7.68	3.11	3.04	1.16
1988	6.34	2.68	11.02	4.61	–4.68	–1.93
1989	–3.06	–1.40	4.87	2.22	–7.93	–3.62
1990	4.27	1.87	–0.20	–0.09	4.46	1.96
1991	8.02	3.39	19.12	7.74	–11.11	–4.35
1992	14.08	4.99	2.17	0.77	11.92	4.22
1993	3.14	1.33	1.19	0.50	1.96	0.83
1994	–0.71	–0.31	–7.15	–3.11	6.44	2.80
1995	14.36	6.25	16.15	7.11	–1.79	–0.86
1996	12.33	5.51	1.75	0.77	10.59	4.74
1997	2.83	1.29	–3.44	–1.54	6.27	2.83
1998	–9.87	–5.12	–14.55	–7.41	4.69	2.28
1999	13.82	6.82	18.48	9.27	–4.66	–2.46
2000	7.43	3.47	11.56	5.52	–4.13	–2.05
2001	–3.22	–1.69	–0.39	–0.21	–2.82	–1.48
2002	–1.84	–0.94	6.27	3.20	–8.11	–4.14
2003	24.35	11.46	2.61	1.11	21.75	10.35
2004	8.94	3.84	1.94	0.76	7.01	3.07
2005	6.57	3.12	2.76	1.32	3.81	1.80
2006	7.17	3.44	6.64	3.11	0.53	0.33
2007	4.50	2.46	–2.64	–1.49	7.14	3.96
Mean	8.33***	3.24***	5.80***	2.26***	2.53**	0.99*
(*t*)	(6.22)	(5.75)	(4.26)	(3.83)	(2.13)	(1.89)
Median	8.00***	3.41***	5.57***	2.24***	3.15**	1.14*
(signed rank)	(232)	(223)	(194)	(189)	(105)	(93)
No. of positive	27/32***	27/32***	26/32***	26/32***	22/32**	22/32**
(*Z*)	(3.89)	(3.89)	(3.54)	(3.54)	(2.12)	(2.12)

Note to Table 3.6:

This table compares returns (in percent) to hedge portfolios, in which equal-weighted portfolios are formed based on F-Score. F-Score represents a summary measure of a firm's financial condition introduced by Piotroski (2000). See Table 3.2 for more details on F-Score. Strong (Weak) represents firms with F-Scores of 5 or greater (0 through 4) and All represents all firms. In Analysis 1, following Piotroski, returns of individual observations are accumulated starting six months after the fiscal year-end, whereas returns under Analysis 2 are accumulated with a common starting date (i.e., July 1 each year). Differences in mean and median are tested with two-sample *t*-test and signed rank test, respectively. *, **, *** denote two-tail significance at the 10%, 5%, and 1% level, respectively.

Table 3.7: One-year market-adjusted returns to F-Score-based value-weighted portfolios

	Returns to Weak, Strong, and All portfolios under Analysis 2					
	Strong portfolios		Weak portfolios		All portfolios	
Year	n	one year market-adjusted return	n	one year market-adjusted return	n	one year market-adjusted return
1976	502	8.09%	125	22.11%	627	9.22%
1977	427	3.79	188	−6.01	615	1.67
1978	416	−4.92	194	−9.35	610	−6.37
1979	410	9.94	204	3.91	614	7.87
1980	433	9.89	195	4.94	628	8.38
1981	462	7.52	191	15.32	653	8.97
1982	397	10.62	256	19.77	653	15.34
1983	480	5.49	191	−6.45	671	2.47
1984	465	−2.71	263	−25.97	728	−9.81
1985	400	26.23	343	5.48	743	17.16
1986	390	7.29	340	10.60	730	8.32
1987	409	14.37	278	−2.34	687	6.86
1988	389	−12.34	280	−8.08	669	−11.02
1989	356	−5.03	298	−10.95	654	−7.08
1990	349	21.75	298	4.72	647	13.94
1991	406	12.77	276	0.35	682	7.56
1992	494	6.63	272	−5.93	766	1.62
1993	389	−0.09	284	−8.70	673	−3.71
1994	383	−8.21	294	−10.71	677	−8.94
1995	440	−8.91	346	−12.84	786	−10.07
1996	522	−2.24	408	−0.11	930	−1.50
1997	540	−10.47	438	−13.44	978	−11.33
1998	444	−22.16	460	11.90	904	−7.70
1999	425	46.01	428	41.13	853	43.71
2000	432	15.05	395	−7.78	827	0.52
2001	411	−8.67	453	6.61	864	−1.01
2002	439	36.62	457	22.56	896	28.32
2003	468	13.40	349	5.11	817	11.63
2004	443	4.21	288	7.07	731	4.72
2005	366	−0.29	337	25.90	703	7.45
2006	322	−0.44	284	8.81	606	3.06
2007	216	−6.42	280	−0.45	496	−3.91
Mean		5.21**		2.72		3.95*
(*t*)		(2.09)		(1.11)		(1.89)
Median		4.85*		2.13		2.77
(signed rank)		(97)		(33)		(84)
No. of positive		18/32		17/32		20/32
(*Z*)		(0.71)		(0.35)		(1.41)

Note to Table 3.7:

This table presents one-year market-adjusted returns (in percent) to value-weighted portfolios, where portfolios are formed based on F-Score. The weight assigned to a firm is determined based on its market capitalization at the beginning of return accumulation date (i.e., July 1 each year). F-Score represents a summary measure of a firm's financial condition introduced by Piotroski (2000). See Table 3.2 for more details on F-Score. Strong (Weak) represents firms with F-Scores of 5 or greater (0 through 4) and All represents all firms. Differences in mean and median are tested with two-sample *t*-test and signed rank test, respectively. *, **, *** denote two-tail significance at the 10%, 5%, and 1% level, respectively.

Table 3.8: The impact of the unknown portfolio weights problem and change in portfolio weighting on the profitability of hedge portfolios

Year	Piotroski (2000) replicated Analysis 1 under equal-weighting: Strong – Weak [A]	Strong – All [B]	Unknown portfolio weights problem eliminated and different portfolio weighting applied Analysis 2 under value-weighting: Strong – Weak [A*]	Strong – All [B*]	Impact of unknown portfolio weights problem and change in portfolio weighting: [A] – [A*]	[B] – [B*]
1976	4.58%	0.99%	–14.02%	–1.13%	18.60%	2.12%
1977	6.08	1.80	9.80	2.12	–3.73	–0.32
1978	10.53	3.29	4.43	1.45	6.10	1.85
1979	16.46	5.49	6.03	2.07	10.43	3.42
1980	13.26	4.04	4.95	1.51	8.31	2.53
1981	10.67	3.18	–7.80	–1.44	18.47	4.63
1982	9.32	3.62	–9.15	–4.72	18.47	8.34
1983	24.68	7.22	11.95	3.02	12.73	4.20
1984	14.41	5.06	23.26	7.10	–8.84	–2.03
1985	18.34	8.52	20.75	9.07	–2.41	–0.55
1986	7.99	3.79	–3.31	–1.03	11.30	4.82
1987	10.73	4.27	16.71	7.51	–5.98	–3.25
1988	6.34	2.68	–4.26	–1.32	10.59	4.00
1989	–3.06	–1.40	5.92	2.04	–8.98	–3.44
1990	4.27	1.87	17.03	7.81	–12.76	–5.94
1991	8.02	3.39	12.41	5.20	–4.39	–1.81
1992	14.08	4.99	12.56	5.01	1.52	–0.02
1993	3.14	1.33	8.61	3.62	–5.47	–2.30
1994	–0.71	–0.31	2.50	0.72	–3.21	–1.03
1995	14.36	6.25	3.93	1.15	10.43	5.09
1996	12.33	5.51	–2.13	–0.74	14.46	6.25
1997	2.83	1.29	2.96	0.85	–0.14	0.44
1998	–9.87	–5.12	–34.05	–14.46	24.19	9.34
1999	13.82	6.82	4.88	2.30	8.94	4.52
2000	7.43	3.47	22.84	14.53	–15.41	–11.06
2001	–3.22	–1.69	–15.29	–7.66	12.07	5.97
2002	–1.84	–0.94	14.06	8.30	–15.90	–9.24
2003	24.35	11.46	8.28	1.76	16.07	9.70
2004	8.94	3.84	–2.86	–0.51	11.80	4.35
2005	6.57	3.12	–26.19	–7.74	32.76	10.86
2006	7.17	3.44	–9.25	–3.50	16.42	6.94
2007	4.50	2.46	–5.97	–2.51	10.48	4.98
Mean	8.33***	3.24***	2.49	1.26	5.84***	1.98**
(*t*)	(6.22)	(5.75)	(1.06)	(1.28)	(2.75)	(2.15)
Median	8.00***	3.41***	4.65	1.48*	9.68***	2.98**
(signed rank)	(232)	(223)	(76)	(91)	(135)	(119)
No. of positive	27/32***	27/32***	20/32	20/32	20/32	20/32
(*Z*)	(3.89)	(3.89)	(1.41)	(1.41)	(1.41)	(1.41)

Note to Table 3.8:

This table compares returns (in percent) to hedge portfolios, in which portfolios are formed based on F-Score. F-Score represents a summary measure of a firm's financial condition introduced by Piotroski (2000). See Table 3.2 for more details on F-Score. Strong (Weak) represents firms with F-Scores of 5 or greater (0 through 4) and All represents all firms. In Analysis 1, following Piotroski, returns of individual observations are accumulated starting six months after the fiscal year-end, whereas returns under Analysis 2 are accumulated with a common starting date (i.e., July 1 each year). For value-weighting, the weight assigned to a firm is determined based on its market capitalization at the beginning of the return accumulation date. Differences in mean and median are tested with two-sample *t*-test and signed rank test, respectively. [*], [**], [***] denote two-tail significance at the 10%, 5%, and 1% level, respectively.

Table 3.9: Profitability of implementing F-Score-based hedge portfolios multiple times a year

Panel A: Profitability of implementing F-Score-based hedge portfolios twice a year

Year	V-W Portfolios formed on January 1 [$FYE_{1/31}$ to $FYE_{6/30}$] n Strong:Weak:All	Strong – Weak	Strong – All	V-W Portfolios formed on July 1 [$FYE_{7/31}$ to $FYE_{12/31}$] n Strong:Weak:All	Strong – Weak	Strong – All
1976	135:31:166	25.42%	6.41%	362:97:459	–22.40%	–1.51%
1977	118:48:166	12.79	2.58	323:132:455	8.94	1.87
1978	116:52:168	–18.42	–4.25	303:141:444	5.42	1.81
1979	121:70:191	5.37	1.88	294:136:430	7.23	2.50
1980	125:64:189	–36.31	–9.56	313:124:437	6.45	1.91
1981	141:66:207	2.08	0.47	330:136:466	–8.55	–1.70
1982	141:66:207	–4.25	–1.63	258:190:448	–9.90	–5.51
1983	133:71:204	14.85	4.56	357:124:481	12.48	3.23
1984	153:88:241	14.57	1.51	321:188:509	20.02	6.80
1985	132:94:226	–7.03	–4.39	272:254:526	20.61	8.55
1986	127:112:239	–17.49	–10.75	274:238:512	–6.58	–2.29
1987	132:113:245	–10.99	–2.61	291:182:473	22.38	10.19
1988	125:105:230	27.27	11.00	266:183:449	–9.02	–3.00
1989	119:106:225	–12.39	–5.12	248:200:448	5.75	1.93
1990	137:88:225	–19.39	–5.08	229:210:439	19.50	9.76
1991	115:120:235	22.73	14.90	296:170:466	16.13	5.94
1992	153:89:242	22.92	9.32	344:187:531	4.75	1.86
1993	136:82:218	3.37	1.31	260:202:462	6.93	3.38
1994	117:88:205	–35.48	–19.50	271:212:483	1.06	0.25
1995	115:103:218	–6.23	–3.26	324:245:569	2.24	0.61
1996	120:118:238	15.62	8.10	406:287:693	0.97	0.29
1997	160:97:257	10.90	2.34	371:351:722	–0.80	–0.26
1998	129:110:239	0.24	0.05	340:368:708	–39.71	–19.02
1999	102:109:211	20.95	13.64	331:323:654	–4.77	–2.05
2000	103:81:184	17.35	3.84	342:324:666	33.72	24.86
2001	88:99:187	24.22	10.37	327:363:690	4.32	2.11
2002	84:91:175	26.92	20.57	359:379:738	12.61	7.40
2003	100:71:171	3.68	0.79	365:261:626	17.62	2.78
2004	101:59:160	–16.77	–3.55	346:234:580	3.96	0.78
2005	73:70:143	6.96	4.09	284:273:557	–31.13	–7.91
2006	69:69:138	10.35	6.44	256:219:475	–12.39	–4.37
2007	61:65:126	–17.61	–5.32	172:225:397	–1.38	–0.57
Avg.	118:84:202			307:224:531		

	Strong – Weak	Strong – All
Mean	2.70	1.56
(t)	(1.30)	(1.67)
Median	4.54*	1.66*
(signed rank)	(248)	(269)
No. of positive	41/64**	41/64**
(Z)	(2.25)	(2.25)

Table 3.9: Profitability of implementing F-Score-based hedge portfolios multiple times a year (cont.)

Panel B: Profitability of implementing F-Score-based hedge portfolios four times a year

Average n (Strong:Weak:All)

V-W Portfolios formed on January 1 [FYE$_{4/30}$ to FYE$_{6/30}$]	V-W Portfolios formed on April 1 [FYE$_{7/31}$ to FYE$_{9/30}$]	V-W Portfolios formed on July 1 [FYE$_{10/30}$ to FYE$_{12/31}$]	V-W Portfolios formed on October 1 [FYE$_{1/31}$ to FYE$_{3/31}$]
60:43:103	52:38:89	256:187:444	60:42:102

	Strong – Weak	Strong – All
Mean	1.89	0.99
(t)	(0.97)	(1.14)
Median	5.02*	1.39**
(signed rank)	(778)	(879)
No. of positive	80/128***	80/128***
(Z)	(2.83)	(2.83)

Panel C: Profitability of implementing F-Score-based hedge portfolios six times a year

Average n (Strong:Weak:All)

V-W Portfolios formed on January 1 [FYE$_{5/31}$ & FYE$_{6/30}$]	V-W Portfolios formed on May 1 [FYE$_{9/30}$ & FYE$_{10/31}$]	V-W Portfolios formed on September 1 [FYE$_{1/31}$ & FYE$_{2/28}$]
49:36:86	44:33:77	30:21:52

V-W Portfolios formed on March 1 [FYE$_{7/31}$ & FYE$_{8/31}$]	V-W Portfolios formed on July 1 [FYE$_{11/30}$ & FYE$_{12/31}$]	V-W Portfolios formed on November 1 [FYE$_{3/31}$ & FYE$_{4/30}$]
22:14:36	242:177:419	39:27:66

	Strong – Weak	Strong – All
Mean	3.38*	2.20**
(t)	(1.71)	(2.46)
Median	2.36**	0.43**
(signed rank)	(1,534)	(1,635)
No. of positive	109/192*	109/192*
(Z)	(1.88)	(1.88)

Note to Table 3.9:

This table reports returns (in percent) to hedge portfolios, in which F-Score-based value-weighted portfolios are formed multiple times a year. F-Score represents a summary measure of a firm's financial condition introduced by Piotroski (2000). See Table 3.2 for more details on F-Score. Strong (Weak) represents firms with F-Scores of 5 or greater (0 through 4) and All represents all firms. For value-weighting, the weight assigned to a firm is determined based on its market capitalization at the beginning of the return accumulation date. Differences in mean and median are tested with two-sample *t*-test and signed rank test, respectively. *, **, *** denote two-tail significance at the 10%, 5%, and 1% level, respectively.

Table 3.10: The impact of the unknown portfolio weights problem and change in portfolio weighting on the profitability of hedge portfolios in the context of Xie (2001)

	Xie (2001) replicated		Unknown portfolio weights problem eliminated		
Year	n Lowest:Highest	Lowest decile – Highest decile (equal-weighting)	n Lowest:Highest	Lowest decile – Highest decile (equal-weighting)	Lowest decile – Highest decile (value-weighting)
1971	130:130	4.91%	130:130	2.83%	12.72%
1972	165:165	5.02	165:165	2.67	–5.88
1973	213:213	14.99	213:213	24.34	–10.90
1974	288:288	7.13	289:289	1.87	25.92
1975	319:319	6.80	320:320	2.23	16.31
1976	320:320	8.36	319:319	3.88	7.29
1977	311:311	12.64	308:308	8.40	–3.67
1978	300:300	4.33	297:297	2.46	0.99
1979	296:296	8.72	293:293	–2.10	0.55
1980	301:301	2.89	297:297	1.91	10.68
1981	302:302	12.59	298:298	12.97	–11.36
1982	320:320	0.34	316:316	3.51	16.48
1983	327:327	7.74	323:323	1.04	18.37
1984	339:339	6.38	332:332	4.50	30.24
1985	352:352	11.69	344:344	16.48	9.01
1986	345:345	5.81	338:338	5.19	2.83
1987	354:354	5.16	346:346	6.95	–3.04
1988	377:377	–3.30	369:369	–7.13	5.88
1989	373:373	5.59	365:365	4.08	–4.12
1990	370:370	14.23	364:364	10.57	10.31
1991	369:369	17.64	361:361	6.98	10.18
1992	373:373	5.31	369:369	–2.44	–16.59
Avg.	311:311		307:307		
Mean		7.50***		5.05***	5.56**
(t)		(7.16)		(3.55)	(2.16)
Median		6.59***		3.70***	6.59*
(signed rank)		(123.5)		(99.5)	(56.5)
No. of positive		21/22***		19/22***	15/22*
(Z)		(4.26)		(3.41)	(1.71)

Note to Table 3.10:

This table reports size-adjusted returns (in percent) to hedge portfolios originally documented in Xie (2001). Portfolio deciles are formed annually based on the ranking of abnormal accruals estimated from cross-sectional version of Jones' (1991) model. The hedge portfolio is formed by taking a long position in the lowest decile portfolio and a short position in the highest decile portfolio. Differences in mean and median are tested with two-sample t-test and signed rank test, respectively. *, **, *** denote two-tail significance at the 10%, 5%, and 1% level, respectively.

Figure 3.1: Illustration of the unknown portfolio weights problem

As of Oct. 30, year *t* — As of Apr. 30, year *t+1* — time

	F-Score		
	0	9	Total
Jun. 30 FYE firms	13	27	40
Dec. 31 FYE firms	?	?	60
Total position	\$3,000 short	\$3,000 long	\$0
Applicable weight to individual firm	?	?	
Individual position	?	?	\$0

	F-Score		
	0	9	Total
Jun. 30 FYE firms	13	27	40
Dec. 31 FYE firms	12	48	60
Total position	\$3,000 short	\$3,000 long	\$0
Applicable weight To individual firm	1/25	1/75	
Individual position	\$120 short	\$40 long	\$0

Chapter 4: Conclusions

In an informationally efficient market, future price changes cannot be forecasted because the current price already fully incorporates the information available to all market participants. Early capital market studies (e.g., Ball and Brown, 1968; Brown, 1968; Fama, Fisher, Jensen, and Roll, 1969) largely supported this notion of market efficiency but later studies (e.g., Foster, Olsen, and Shevlin, 1984; Bernard and Thomas, 1990; Sloan, 1996) with more refined data challenge market efficiency. This dissertation investigates explanations for stock market anomalies related to accounting information as documented by Dichev (1998) and Piotroski (2000).

The bankruptcy risk anomaly, originally documented by Dichev (1998), refers to the empirical regularity in which a composite measure of bankruptcy risk predicts future returns. Taken as a whole, the first study of the dissertation demonstrates that Dichev's bankruptcy risk anomaly is a manifestation of investors' under (over)-pricing of cash flows (accrual) component of earnings, i.e., Sloan's (1996) accrual anomaly. This study contributes to the existing research in several ways. First, it provides an alternative explanation for Dichev's findings. In this sense, this study is complementary to the explanations provided by Griffin and Lemmon (2002) and George and Hwang (2010). Second, the study highlights that the choice of bankruptcy risk measure can lead to different conclusions. Third, by relating the bankruptcy risk anomaly to the accrual anomaly, it contributes to a stream of research that investigates the relationship between anomalies (e.g., Collins and Hribar, 2000; Desai, Rajgopal, and Venkatachalam, 2004).

One limitation of the first study is that my findings are limited to the case of Dichev (1998). Subsequent studies that follow Dichev use their own measures of bankruptcy risk and thus, this study cannot address their results. As shown in my first study, their findings could be a manifestation of other anomalies but such an assessment would be beyond the scope of this study.

The second study of the dissertation investigates the effects of two potentially problematic research design choices that are often made in accounting-based studies of anomalies. I explore these issues by re-examining the results in Piotroski (2000). I find that the relationship between Piotroski's fundamental signals and subsequent returns is partly driven by the choice of return accumulation periods and the use of equally weighted returns. When the research design controls for both problems, the relationship between his fundamental signals and subsequent returns disappears. Because the methods used in Piotroski are typical of those often employed in the accounting literature, this study suggests that evidence of profitable trading strategies and market inefficiency in the literature is likely to be overstated. Therefore, it contributes to a stream of research that proposes research design choices or flaws as an explanation of detected mispricing (e.g., Lo and MacKinlay, 1990; Kothari, Shanken, and Sloan, 1995; Kothari, Sabino, and Zack, 2005; Kraft, Leone, and Wasley, 2006).

Studies that document anomalies, to varying extent, are subject to criticisms that research design flaws or omitted risk factors are responsible for the anomalous returns. Such studies need to be especially careful to rule out explanations other than informational inefficiency in securities markets, and in some cases, existing anomalies documented in prior research. This dissertation explores alternative explanations for stock market anomalies documented in two separate studies by Dichev (1998) and Piotroski (2000). In the case of Dichev, I find that what appeared initially to be a new anomaly related to bankruptcy risk is a manifestation of the anomaly previously documented by Sloan (1996). In the case of Piotroski, I show that the relationship between his fundamental signals and subsequent returns is an unintended consequence of his research design choices, demonstrating that seemingly innocuous research design choices can have substantial consequences when exploring evidence of stock market inefficiency.

References

Alford, A. W., J. J. Jones, and M. E. Zmijewski. 1994. Extensions and violations of the statutory SEC form 10-K filing requirements. *Journal of Accounting and Economics* 17 (1–2): 229–254.

Ali, A., L. S. Hwang, and M. A. Trombley. 2003. Residual-income-based valuation predicts future stock returns: evidence on mispricing vs. risk explanations. *The Accounting Review* 78 (2): 377–396.

Altman, E. I. 1968. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The journal of Finance* 23 (4): 589–609.

American Institute of Certified Public Accountant (AICPA). 1963. Accounting Principle Board Opinion No. 3: *The Statement of Source and Application of Funds*. New York: AICPA.

Anginer, D. and C. Yildizhan. 2010. Is there a distress risk anomaly? Corporate bond spread as a proxy for default risk. Working paper, University of Michigan.

Ball, R. and P. Brown. 1968. An empirical evaluation of accounting income numbers. *Journal of Accounting Research* 6 (2): 159–177.

Ball, R. and S. P. Kothari. 1989. Nonstationary expected returns: implications for tests of market efficiency and serial correlation in returns. *Journal of Financial Economics* 25 (1): 51–74.

Barth, M. E., W. H. Beaver, and W. R. Landsman. 1992. The market valuation implications of net periodic pension cost components. *Journal of Accounting and Economics* 15 (1) 27–62.

Basu, S. 2004. What do we learn from two new accounting-based stock market anomalies? *Journal of Accounting and Economics* 38 (1): 333–348.

Beaver, W. H. 1966. Financial ratios as predictors of failure. *Journal of Accounting Research* 4 (Supplement): 71–111.

Beaver, W. H. 1968. The information content of annual earnings announcement. *Journal of Accounting Research* 6 (Supplement): 67–92.

Beaver, W. H., C. Eger, S. Ryan, and M. Wolfson. 1989. Financial reporting, supplemental disclosure, and bank share prices. *Journal of Accounting Research* 27 (2): 157–178.

Begley, J., J. Ming, and S. Watts. 1996. Bankruptcy classification errors in the 1980s: An empirical analysis of Altman's and Ohlson's models. *Review of Accounting Studies* 1 (4): 267–284.

Beneish, M. D. and M. E. Vargus. 2002. Insider trading, earnings quality, and accrual mispricing. *The Accounting Review* 77 (4): 755–791.

Bernard, V. L. and J. K. Thomas. 1989. Post-earnings-announcement drift: delayed price response or risk premium? *Journal of Accounting Research* 27 (Supplement): 1–48.

Bernard, V. L. and J. K. Thomas. 1990. Evidence that stock prices do not fully reflect the implication of current earnings for future earnings. *Journal of Accounting and Economics* 13 (4): 305–340.

Bhushan, R. 1994. An informational efficiency perspective on the post-earnings announcement drift. *Journal of Accounting and Economics* 18 (1): 45–65.

Black, F. and M. Scholes. 1973. The pricing of options and corporate liabilities. *Journal of Political Economy* 81 (3): 637–654.

Boehme, R. D., B. R. Danielsen, S. M. Sorescu. 2006. Short-sale constraints, differences of opinion, and overvaluation. *Journal of Financial and Quantitative Analysis* 41 (2): 455–487.

Campbell, J. Y., J. Hilscher, and J. Szilagyi. 2008. In search of distress risk. *The Journal of Finance* 63 (6): 2899–2939.

Campbell, J. Y. and T. Vuolteenaho. 2004. Bad beta, good beta. *The American Economic Review* 94 (5): 1249–1275.

Carhart, M. M. 1997. On persistence in mutual fund performance. *The Journal of Finance* 52 (1): 57–82.

Chan, K. C. 1988. On the contrarian investment strategy. *Journal of Business* 61 (2): 147–163.

Chava, S. and A. Purnanandam. 2010. Is default risk negatively related to stock returns? *The Review of Financial Studies* 23 (6): 2523–2559.

Collins, D. W. and P. Hribar. 2000. Earnings-based and accrual-based market anomalies: One effect or two? *Journal of Accounting and Economics* 29 (1): 101–123.

Dechow, P. M., S. A. Richardson, and R. G. Sloan. 2008. The persistence and pricing of the cash component of earnings. *Journal of Accounting Research* 46 (3): 537–566.

DeFond, M. L., J. Jiambalvo. 1994. Debt covenant violation and manipulation of accruals. *Journal of Accounting and Economics* 17 (1&2): 145–176.

Desai, H., S. Rajgopal, and M. Venkatachalam. 2004. Value-glamour and accruals mispricing: One anomaly or two? *The Accounting Review* 79 (2): 355–385.

Dichev, I. D. 1998. Is the risk of bankruptcy a systematic risk? *The Journal of Finance* 53 (3): 1131–1147.

Fairfield, P. M., J. S. Whisenant, and T. L. Yohn. 2003. Accrued earnings and growth: implications for future profitability and market mispricing. *The Accounting Review* 78 (1): 353–371.

Fama, E. F. 1970. Efficient capital markets: a review of theory and empirical work. *The Journal of Finance* 25 (2): 383–417.

Fama, E. F. 1998. Market efficiency, long-term returns, and behavioral finance. *Journal of Financial Economics* 49 (3): 283–306.

Fama, E. F., L. Fisher, M. C. Jensen, and R. Roll. 1969. The adjustment of stock prices to new information. *International Economic Review* 10 (1): 1–21.

Fama, E. F. and K. R. French. 1992. The cross-section of expected stock returns. *The Journal of Finance* 47 (2): 427–465.

Fama, E. F. and K. R. French. 2008. Dissecting anomalies. *The Journal of Finance* 63 (3): 1653–1678.

Foster, G, C. Olsen, and T. Shevlin. 1984. Earnings release, anomalies and the behavior of security returns. *The Accounting Review* 59 (4): 574–603.

French, K. R., G. W. Schwert, and R. F. Stambaugh. 1987. Expected stock returns and stock market volatility. *Journal of Financial Economics* 19 (1): 3–30.

George, T. J. and C. Y. Hwang. 2010. A resolution of the distress risk and leverage puzzles in the cross section of stock returns. *Journal of Financial Economics* 96 (1): 56–79.

Givoly, D. and C. Hayn. 2000. The changing time-series properties of earnings, cash flows and accruals: Has financial reporting become more conservative? *Journal of Accounting and Economics* 29 (3): 287–320.

Griffin, J. M. and M. L. Lemmon. 2002. Book-to-market equity, distress risk, and stock returns. *The Journal of Finance* 57 (5): 2317–2336.

Guay, W. 2000. Discussion of value investing: the use of historical financial statement information to separate winners from losers. *Journal of Accounting Research* 38 (Supplement): 43–51.

Hillegeist, S., E. Keating, D. Cram, and K. Lundstedt. 2004. Assessing the probability of bankruptcy. *Review of Accounting Studies* 9 (1): 5–34.

Hirshleifer, D., K. Hou, S. H. Teoh, and Y. Zhang. 2004. Do investors overvalue firms with bloated balance sheet? *Journal of Accounting and Economics* 38 (Supplement): 297−331.

Hirshleifer, D., J. N. James, L, A. Myers, and S. H. Teoh. 2008. Do individual investors cause post-earnings announcement drift? Direct evidence from personal trades. *The Accounting Review* 83 (6): 1521−1550.

Hirshleifer, D., S. S. Lim, and S. H. Teoh. 2009. Driven to distraction: Extraneous events and underreaction to earnings news. *The Journal of Finance* 64 (5): 2287−2323.

Hirshleifer, D., S. S. Lim, and S. H. Teoh. 2011. Limited investor attention and stock market misreactions to accounting information. *Review of Asset Pricing Studies* 1 (1): 35−73.

Hirshleifer, D. and S. H. Teoh. 2003. Limited attention, information disclosure, and financial reporting. *Journal of Accounting and Economics* 36 (1−3): 337−386.

Houge, T. and T. Loughran. 2000. Cash flow is king? Cognitive errors by investors. *Journal of Psychology & Financial Markets* 1 (3&4): 161−175.

Hribar, P. and D. W. Collins. 2002. Errors in estimating accruals: Implications for empirical research. *Journal of Accounting Research* 40 (1): 105−134.

Jones, C. M. and O. A. Lamont. 2002. Short-sale constraints and stock returns. *Journal of Financial Economics* 66 (2−3): 207−239.

Jones, J. J. 1991. Earnings management during import relief investigations. *Journal of Accounting Research* 29 (2): 193−228.

Keim, D. B. 2008. Financial market anomalies, in Durlauf, S. N. and L. E. Blume (eds.). *The New Palgrave Dictionary of Economics.* New York, NY: Palgrave Macmillian

Keim, D. B. and A. Madhavan. 1997. Transaction costs and investment style: an inter-exchange analysis of institutional equity trades. *Journal of Financial Economics* 46 (3): 265−292.

Khan, M. 2008. Are accruals mispriced? Evidence from tests of an Intertemporal Capital Asset Pricing Model. *Journal of Accounting and Economics* 45 (1): 55−77.

Kim, D. and M. Kim. 2003. A multifactor explanation of post-earnings announcement drift. *Journal of Financial and Quantitative Analysis* 38 (2): 383−398.

Kothari, S. P. 2001. Capital markets research in accounting. *Journal of Accounting and Economics* 31 (1−3): 105−231.

Kothari, S. P., J. S. Sabino, and T. Zach. 2005. Implications of survival and data trimming for tests of market efficiency. *Journal of Accounting and Economics* 39 (1): 129−161.

Kothari, S. P., J. Shanken, and R. G. Sloan. 1995. Another look at the cross-section of expected returns. *The Journal of Finance* 50 (1): 185−224.

Kraft, A., A. J. Leone, and C. Wasley. 2006. An analysis of the theories and explanations offered for the mispricing of accruals and accrual components. *Journal of Accounting Research* 44 (2): 297−339.

Lakonishok, J., A. Shleifer, and R. W. Vishny. 1994. Contrarian investment, extrapolation, and risk. *The Journal of Finance* 49 (5): 1541−1578.

Lamont, O. A. and R. H. Thaler. 2003. Can the market add and subtract? Mispricing in tech stock carve-outs. *Journal of Political Economy* 111 (2): 227−268.

Landsman, W. R. and E. L. Maydew. 1999. Beaver (1968) revisited: has the information content of annual earnings announcements declined in the past three decades? Working paper, University of North Carolina.

Lesmond, D. A., J. P. Ogden, and C. A. Trzcinka. 1999. A new estimate of transaction costs. *The Review of Financial Studies* 12 (5): 1113−1141.

Lewellen, J. 2010. Accounting anomalies and fundamental analysis: an alternative view. *Journal of Accounting and Economics* 50 (2&3): 455−466.

Lo, A. W. and A. C. MacKinlay. 1990. Data-snooping biases in tests of financial asset pricing models. *The Review of Financial Studies* 3 (3): 431−467.

Malkiel, B. G. 1995. Returns from investing in equity mutual funds 1971 to 1991. *The Journal of Finance* 50 (2): 549−572.

Mashruwala, C., S. Rajgopal, and T. Shevlin. 2006. Why is the accrual anomaly not arbitraged away? The role of idiosyncratic risk and transaction costs. *Journal of Accounting and Economics* 42 (1): 3−33.

McDonald, R. L. 2002. *Derivative Markets*. Boston, MA: Addison-Wesley.

Mendenhall, R. R. 2004. Arbitrage risk and post-earnings-announcement drift. *Journal of Business* 77 (4): 875−894.

Merton, R. C. 1974. On the pricing of corporate debt: The risk structure of interest rates. *The Journal of Finance* 29 (2): 449−470.

Miller, M. H. and K. Rock. 1985. Dividend policy under asymmetric information. *The Journal of Finance* 40 (4): 1031−1051.

Myers, S. C. and N. S. Majluf. 1984. Corporate financing and investment decision when firms have information that investors do not have. *Journal of Financial Economics* 13 (2): 187−221.

Narayanamoorthy, G. 2006. Conservatism and cross-sectional variation in the post-earnings announcement drift. *Journal of Accounting Research* 44 (4): 763−789.

Ng, J., T. O. Rusticus, andR. S. Verdi. 2008. Implications of transaction costs for the post-earnings announcement drift. *Journal of Accounting Research* 46 (3): 661−696.

Ohlson, J. A. 1980. Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research* 18 (1): 109−131.

Penman, S. H., S. A. Richardson, and I. Tuna. 2007. The book-to-price effect in stock returns: accounting for leverage. *Journal of Accounting Research* 45 (2): 427−467.

Penman, S. H. and X. Zhang. 2002. Accounting conservatism, the quality of earnings, and stock returns. *The Accounting Review* 77 (2): 237−264.

Piotroski, J. D. 2000. Value investing: The use of historical financial statement information to separate winners from losers. *Journal of Accounting Research* 38 (Supplement): 1−41.

Pincus, M., S. Rajgopal, and M. Venkatachalam. 2007. The accrual anomaly: international evidence. *The Accounting Review* 82 (1): 169−203.

Richardson, S., R. G. Sloan, M. T. Soliman, and I. Tuna. 2005. Accrual reliability, earnings persistence and stock prices. *Journal of Accounting and Economics* 39 (3): 437−485.

Richardson, S., I. Tuna, P. Wysocki. 2010. Accounting anomalies and fundamental analysis: a review of recent research advances. *Journal of Accounting and Economics* 50 (2&3): 410−454.

Ritter, J. R. and I. Welch. 2002. A review of IPO activity, pricing, and allocations. *The Journal of Finance* 57 (4): 1795−1828.

Rosenberg, B., K. Reid, and R. Lanstein. 1984. Persuasive evidence of market inefficiency. *Journal of Portfolio Management* 11 (3): 9−16.

Shumway, T. 2001. Forecasting bankruptcy more accurately: A simple hazard model. *Journal of Business* 74 (1): 101−124.

Sloan, R. G. 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review* 71 (3): 289−315.

Subramanyam, K. R. 1996. The pricing of discretionary accruals. *Journal of Accounting and Economics* 22 (1−3): 249−281.

Vassalou, M. and Y. Xing. 2004. Default risk in equity returns. *The Journal of Finance* 59 (2): 831–868.

Xie, H. 2001. The mispricing of abnormal accrual. *The Accounting Review* 76 (3): 357–374.

Zhang, X. F. 2007. Accruals, investment, and the accrual anomaly. *The Accounting Review* 82 (5): 1333–1363.

Zmijewski, M. E. 1984. Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research* 22 (Supplement): 59–82.

Appendices

Appendix A Estimation of Hillegeist et al.'s (2004) *BSM-Prob*

Hillegeist et al. (2004) modify the original option-pricing equation as follows:

$$V_E = V_A e^{-\delta T} N(d_1) - X e^{-rT} N(d_2) + (1 - e^{-\delta T}) V_A, \tag{A1}$$

where $N(d_1)$ and $N(d_2)$ are the standard cumulative normal distribution of d_1 and d_2, respectively, and

$$d_1 = \frac{\ln[V_A / X] + (r - \delta + (\sigma_A^2 / 2))T}{\sigma_A \sqrt{T}}, \tag{A2}$$

and

$$d_2 = d_1 - \sigma_A \sqrt{T}, \tag{A3}$$

V_E is the current market value of equity; V_A is the current market value of assets; X is the face value of debt maturing at time T; r is the continuously compounded risk-free rate; δ is the dividend rate expressed in terms of V_A; and σ_A is the standard deviation of assets returns. They make two modifications to the traditional option-pricing model. The $V_A e^{-\delta T}$ term accounts for the reduction in the value of the assets due to the dividends that are distributed before time T. The addition of the $(1 - e^{-\delta T})V_A$ term is necessary because it is the equity holders who receive the dividends.

Under the BSM model, the probability of bankruptcy is simply the probability that the market value of assets, V_A, is less than the face value of the liabilities, X, at time T. The BSM model assumes that the natural log of future asset values is distributed normally as follows:

$$\ln V_A(t) \sim N\left[\ln V_A + \left(\mu - \delta - \frac{\sigma_A^2}{2}\right)t, \sigma_A^2 t\right], \tag{A4}$$

where μ is the continuously compounded expected return on assets. Then, as shown in McDonald (2002, p. 604), the probability that $V_A(T) < X$ is, as follows:

$$N\left(-\frac{\ln(V_A / X)+(\mu-\delta-(\sigma_A^2 / 2))T}{\sigma_A\sqrt{T}}\right)=BSM\text{-}Prob. \tag{A5}$$

Equation (A5) shows that the probability of bankruptcy is a function of the distance between the current value of the firm's assets and the face value of its liabilities $\ln(V_A/X)$ adjusted for the expected growth in asset values ($\mu - \delta$) relative to asset volatility (σ_A). Note that the value of the call option in equation (A1) is not a function of μ because equation (A1) is derived under the assumption of risk-neutrality. However, the probability of bankruptcy depends upon the actual distribution of future asset values, which is a function of μ.

To empirically estimate *BSM-Prob* from equation (A5), the market value of assets, V_A, asset volatility, σ_A, and the expected return on assets, μ, need to be estimated because these values are not directly observable. In the first step, Hillegeist et al. (2004) estimate the values of V_A and σ_A by simultaneously solving equation (A1) and the following optimal hedge equation:[28]

$$N(d_1)=\frac{V_E\sigma_E}{V_A e^{-\delta T}\sigma_A}. \tag{A6}$$

V_E is set equal to the total market value of equity at the end of the firm's fiscal year ending in calendar year $t-1$. σ_E is computed using daily return data from the CRSP over the entire fiscal year ending in calendar year $t-1$. The strike price X is set equal to the book value of total liabilities, T equals one year, and r is the one-year Treasury bill rate. The dividend rate, δ, is the sum of common and preferred dividends divided by sum of market value of equity plus total liabilities at

[28] There are cases where the iterative process can not solve the two equations. This occurs because of the boundary condition which requires solutions of the equations to be positive. Accordingly, few observations with valid data are lost; the number of lost observations is less than 0.01 percent of the total number of firm-years in my sample. Thus, I do not expect that this deletion will make any change in the tenor of the results.

the end of the fiscal year. Note that all applicable accounting data are from financial statements with fiscal years ending calendar year *t−1* so that *BSM-Prob* is estimated on an *ex ante* basis.

In the second step, the expected market return on assets, μ, is estimated based on the actual return on assets during the previous year as follows:

$$\mu(t) = \frac{V_A(t) + dividends - V_A(t-1)}{V_A(t-1)}, \tag{A7}$$

where *dividends* is the sum of the common and preferred dividends declared during the year. Note that this process is based on the estimates of V_A that were computed in the first step. When expected returns are below the risk-free rate (above one), Hillegeist et al. (2004) set them equal to the risk-free rate (one).

Appendix B An association of a variable with a score: An econometric interpretation

Without a loss of generality, suppose that there are n firms and each firm has its score S, which is a linear combination of k components (e.g., accounting variables). Using matrices, I can express scores for n firms as:

$$\mathbf{S} = \mathbf{X}\boldsymbol{\alpha}, \tag{B1}$$

where $\mathbf{S} = n \times 1$, $\mathbf{X} = n \times k$, and $\boldsymbol{\alpha} = k \times 1$. Then, an association of a variable Y (e.g., returns) with the score S can be written as:

$$\beta = (\mathbf{S}'\mathbf{S})^{-1}\mathbf{S}'\mathbf{Y} \neq 0, \tag{B2}$$

where $\mathbf{Y}$ is an $n \times 1$. Substituting (B1) into (B2), I obtain:

$$\beta = (\boldsymbol{\alpha}'\mathbf{X}'\mathbf{X}\boldsymbol{\alpha})^{-1}\boldsymbol{\alpha}'\mathbf{X}'\mathbf{Y} \neq 0. \tag{B3}$$

With respect to the numerator ($\boldsymbol{\alpha}'\mathbf{X}'\mathbf{Y}$), since $\boldsymbol{\alpha} \neq \mathbf{0}$, the following holds:

$$\mathbf{X}'\mathbf{Y} \neq \mathbf{0}. \tag{B4}$$

The denominator ($\boldsymbol{\alpha}'\mathbf{X}'\mathbf{X}\boldsymbol{\alpha}$) is non-zero as long as $\mathbf{X}'\mathbf{X}$ is positive definite (i.e., rank ($\mathbf{X}$) = k). This condition is satisfied as long as one component of the score is not a linear combination of the other components. Thus, it follows that, in a regression of variable Y on X (the component of the score), the resulting coefficients is a non-zero vector:

$$(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} = \boldsymbol{\gamma} = \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_k \end{bmatrix} \neq \mathbf{0}. \tag{B5}$$

Therefore, when there is an association between a variable and a score, it must be true that one or more components of the score have nontrivial associations with the variable, holding all other components constant.