Rev Account Stud (2012) 17:969–1010
DOI 10.1007/s11142-012-9186-7

Do differences in financial reporting attributes impair the predictive ability of financial ratios for bankruptcy?

William H. Beaver · Maria Correia · Maureen F. McNichols

Published online: 3 April 2012
© Springer Science+Business Media, LLC 2012

Abstract This study explores the effect of cross-sectional and time-series differences in financial reporting attributes on the predictive ability of financial ratios for bankruptcy. We identify proxies for discretion over financial reporting, the importance of intangible assets, the comprehensiveness of the accounting model and recognition of losses. Each of our proxies for financial reporting attributes is associated with financial ratios that are less informative in predicting bankruptcy. Furthermore, our time-series tests reveal a decline in the predictive ability of financial ratios for bankruptcy and document that this decline is associated with our measures of financial reporting attributes.

Keywords Bankruptcy · Accounting information · Financial ratios

JEL Classification M41 · G14 · G33 · C41

1 Introduction

This paper examines whether the informativeness of financial ratios for bankruptcy prediction varies with attributes hypothesized in the accounting literature to influence financial reporting quality. These attributes include management's exercise of discretion over financial reporting, the importance of intangibles, the

W. H. Beaver · M. F. McNichols (✉)
Graduate School of Business, Stanford University, Stanford, CA 94305, USA
e-mail: fmcnich@stanford.edu

W. H. Beaver
e-mail: fbeaver@stanford.edu

M. Correia
London Business School, London NW1 4SA, UK
e-mail: mcorreia@london.edu

comprehensiveness of the financial reporting model, and the reporting of losses. Collectively, these attributes present some of the most significant challenges to the financial reporting model and its ability to reflect information about firm performance and condition. Our study examines whether variation in these attributes in cross-section and over time is associated with the predictive ability of financial ratios for bankruptcy.

This question is of interest for several reasons. Beginning with Altman (1968) and Beaver (1965, 1966), researchers have found that accounting-based models have significant explanatory power for bankruptcy. While bankruptcy prediction is not the sole purpose of financial reporting, the information imbedded in accounting numbers about the likelihood of bankruptcy serves a role distinct from the informativeness of accounting numbers for security returns or for other purposes. Given the wide use of bankruptcy prediction models in practice and research and the significance of bankruptcy to investors and lenders, the informativeness of financial ratios for bankruptcy prediction is of interest in its own right. Our analyses extend prior literature by exploring how accounting characteristics related to financial statement quality affect the informativeness of accounting numbers for bankruptcy prediction. Our findings bear on how financial reporting qualities affect the informativeness of accounting numbers for an important prediction task and allow an assessment of the significance of changes over time and of cross-sectional differences from the vantage point of predictive ability for bankruptcy.

In addition, our study contributes to the literature by examining three model forms—accounting-based, market-based, and combined. This permits us to compare the predictive ability of each class of explanatory variables. This is essential for our research design because we are studying the effect of characteristics of the financial statements on predictive power. It is important for us to know whether accounting-based and market-based variables are differentially affected and also how the combined model performs. By comparing the performance of the accounting model to the performance of the market model, we have a benchmark that controls for potential differences in the degree of uncertainty inherent in bankruptcy prediction.

We view the accounting-based variables as reflecting a different set of information than the market, which in addition reflects other available information. Accounting numbers are a proper subset of all information potentially reflected in market prices. As a consequence, it may well be that characteristics that reduce the predictive power of accounting-based variables have less effect on the predictive power of market-based variables. On the other hand, the predictive power of the market-based variables may also be undermined by the presence of these characteristics. In the latter case, there are a number of interpretations, which we will discuss in detail.

We find that each of our proxies for the exercise of discretion in financial reporting—that is, existence of a restatement and discretionary accruals—is associated with a significant deterioration in the predictive power of the financial ratio-based model. In addition, the presence of discretion impairs the predictive ability of not only the accounting-based model but also the market-based and combined model. In other words, the total mix of information reflected in market-based variables, of which accounting data are a subset, does not offset or compensate for the effects of discretion.

We also find that the presence of intangible assets, as measured by research and development intensity, has a systematic effect on predictive ability. In particular, the predictive power of the accounting-based model is lower for firms with a high degree of research and development intensity. This is consistent with the concern raised by Lev and Zarowin (1999) that accounting for intangibles results in less informative financial statements and with the findings of Franzen et al. (2007).

We find that the predictive power of the bankruptcy model varies with our proxy for the comprehensiveness of the accounting model—how close the book-to-market ratio is to one—in ways that suggest its effect is nonlinear. Specifically, those firm-years with low to medium positive book-to-market ratios are most informative, consistent with more informative financial statements when the book value of equity is closer to the market value of equity. Those firm-years with high book-to-market ratios are next most informative, and the financial ratios of firms with negative book-to-market ratios are least informative. In other words, when financial statements fail to recognize asset or liability values or both, the predictive ability of financial ratios is impaired.

We find that the incurrence of a loss significantly increases the conditional probability of bankruptcy. However, we also find that the predictive power of the bankruptcy model for loss firm-years is lower than for nonloss firm-years because of deterioration in the incremental explanatory power of the remaining variables.

Finally, we conduct time-series tests to assess whether the effects of financial reporting attributes on predictive ability observed in the cross-section have implications for the predictive ability of financial ratios for bankruptcy over time. We find that there is a significant time trend in the frequency of restatements, larger magnitudes of discretionary accruals, greater R&D intensity, book-to-market ratios that are further from one, and losses. In addition, we find that these variables are individually significant in explaining differences in predictive ability over time. Because these variables are highly correlated, however, it is difficult to isolate individual, incremental effects.

Although the market model generally exhibited lower predictive power than the accounting model in the cross-sectional analysis, the market model exhibits no declining time trend and differences in its predictive power over time are uncorrelated with our partitioning variables. These findings suggest that the changes in financial reporting attributes we document contribute to less informative financial ratios, as assessed by bankruptcy prediction. Furthermore, we find that the combined model exhibits a declining time trend in predictive power and that this is associated with our partitioning variables. These findings indicate that the market variables included in the market and combined models do not fully compensate for the loss of information over time.

The paper proceeds as follows. Section 2 discusses the prior literature. Section 3 discusses our hypotheses. Section 4 discusses the estimation models. Section 5 presents sample properties, measurement of the variables and descriptive statistics. Section 6 discusses the results, and Sect. 7 concludes.

2 Prior literature

A large literature in accounting examines whether the informativeness of financial statements has declined over time (Brown et al. 1999; Francis and Schipper 1999; Francis et al. 2002; Landsman and Maydew 2002). This literature has almost exclusively examined this issue in the context of explaining security returns. In contrast, our dependent variable of interest is bankruptcy. The ability to predict bankruptcy represents a different use of accounting data from prior research and is of interest in its own right. A helpful feature of our research approach is that we can compare the predictive ability of financial ratios with the predictive ability of market-related information over time. Our findings are thus informative to those interested in assessing bankruptcy risk and in understanding whether certain financial reporting attributes are associated with less informative financial ratios for bankruptcy prediction.

As such, our findings are relevant to the literature in accounting and finance on bankruptcy prediction. Recent contributions to this literature include those of Shumway (2001); Chava and Jarrow (2004); Beaver et al. (2005); Franzen et al. (2007); and Campbell et al. (2008). Shumway proposes a hazard model based on accounting and market variables that produces consistent and accurate estimates of the likelihood of bankruptcy. Chava and Jarrow (2004) examine the role of industry effects in a model with accounting and market variables. Beaver et al. (2005) examine whether there have been changes from 1962 to 2002 in the ability of financial ratios to predict bankruptcy and find only a slight decline. Franzen et al. (2007) examine the effect of R&D intensity on the predictive ability of accounting-based bankruptcy models. Campbell et al. (2008) begin with a model of distress risk that incorporates accounting and market variables similar to those used by Shumway and consider alternative measures and additional variables, including Moody's KMV measure of distance to default. They then use their default risk measure to test whether there is a risk premium embedded in security returns incremental to size and value factors.

Our study differs from these studies in several important respects. We examine three model forms—accounting-based, market-based, and combined. This permits us to compare predictive ability of each class of explanatory variables. We view the accounting-based variables as potentially reflecting a subset of information to the market, which, together with other available information, affects the market-based variables used in bankruptcy prediction models. Because market-based prediction models potentially reflect a much richer set of information than financial ratios, attributes that adversely affect accounting-based prediction models may have a different effect on market-based prediction models.

Our study also differs in that our main purpose is to examine the effect of financial reporting attributes on predictive power, which is not examined by Shumway (2001), Chava and Jarrow (2004), Beaver et al. (2005), Campbell et al. (2008), and Bharath and Shumway (2008). Also, Campbell et al. (2008) examine whether default risk can explain some of the return anomalies, which is beyond the scope of our study.

In particular, our study differs in key respects from Franzen et al. (2007). They focus on the effect of expensing research and development on the predictive ability of financial ratios used in bankruptcy prediction. We take a broader view, examining several proxies for financial statement characteristics, including restatements, discretionary accruals, incurrence of an accounting loss, and the market-to-book ratio, in addition to research and development expenditures. In addition, Franzen et al. does not examine the relative performance of accounting-based predictions relative to market-based predictions. This is an important aspect of our study because accounting numbers and market-price-based variables potentially reflect different information and may be differentially affected by the accounting characteristics.

Our study also differs in four key respects from Beaver et al. (2005), who test whether financial ratios have lower predictive ability for bankruptcy in 1994 through 2002 relative to 1962 through 1993. First, we directly examine the relation between the predictive ability of financial ratios for bankruptcy and measures of the influence of discretion on financial statements, the intensity of intangibles, the comprehensiveness of financial statements as reflected in book-to-market ratios, and loss recognition. Beaver et al. (2005) examine a time-series trend in predictive ability and offer no evidence that it is in fact due to accounting characteristics. In contrast, we explicitly measure several proxies for accounting characteristics, examine the effect of these attributes in a cross-sectional research design, and test for their effect on differences in predictive ability in cross-sectional as well as time-series tests. Third, we consider an expanded sample that includes NASDAQ firms, resulting in greater cross-sectional variation in the financial reporting attributes and more powerful tests of their effects. The power arises because the number of firm-years that are bankruptcy years are approximately 1 % of the total sample. The key feature of our study involves partitioning these bankruptcy firm-years even further by accounting characteristic. Hence, increasing the sample size which increases the number of bankruptcy firm-years represents a potentially significant increase in power of the tests. Fourth, we find significant differences in predictive ability in the cross-section and over time that are associated with our proxies for financial reporting attributes. The differences in our inferences suggest that the direct measure of cross-sectional differences in financial reporting attributes has resulted in a more powerful design, presumably because the cross-sectional differences in these attributes are large relative to differences over time.

3 Hypotheses

3.1 Effects of discretion

Academic research has examined the presence of discretion in financial reporting extensively.[1] Managers can exercise discretion in the financial statements opportunistically or to improve the informativeness of financial statements. Prior literature documents a number of settings in which management aims to obscure the underlying financial condition of the firm opportunistically. The incentives for

[1] McNichols (2000) and Dechow and Schrand (2004) provide reviews of this literature.

misreporting include influencing security price, lowering costs of equity and debt, increasing compensation for management, deterring actions of creditors, and reducing the probability of management removal. Watts and Zimmerman (1990), McNichols (2000), and Beaver (2002), among others, discuss these motivations in more detail. In the second scenario, suggested by the signaling literature, management exercises discretion over its financial statements to signal its private information about the firm. There is some evidence in favor of the signaling hypothesis in the banking industry with respect to loan loss provisions (Beaver and Engel 1996; Wahlen 1994). Moreover, to the extent that both signaling and opportunistic behavior are present in the data, the informativeness of financial statements could be impaired, enhanced, or unchanged overall. From this perspective, the purpose of our study is to understand what the net effect is and how discretion contributes to it.

Our study contributes to the literature on accounting quality by examining the effect of two measures of discretion in financial reporting on the predictive power of financial statements for bankruptcy. Our measures of discretion are the presence or absence of a subsequent restatement of financial statements for a firm-year and the magnitude of an estimate of discretionary accruals using the Dechow et al. (1995) model. The null hypothesis is that discretion does not impair predictive ability of financial ratios for bankruptcy. Taking the view that discretion is used predominantly in an opportunistic rather than an informative fashion, the first (alternate) hypothesis is that increased discretion in financial reporting impairs the predictive ability of financial ratios.

Our first proxy for discretion in financial statements is the existence of a violation of GAAP that results in a restatement of the financial statements. FASB statements, SEC enforcement actions, and plaintiffs in securities litigation all assert that violations of GAAP reduce the informativeness of financial statements. However, there is little direct evidence that financial statements that do not comply with GAAP are less informative. The principal conjecture in the literature, as well as by regulators and the professional accounting community, is that the violation of GAAP undermines the informativeness of financial statements. Note that the identity of the restatement firm-years is only known subsequently (for example, possibly as much as several years later). As with any of the accounting characteristic variables, a finding of deteriorated predictive power may be due to that variable or omitted correlated variables. The paper will discuss this caveat further in the discussion of the findings.

Our second proxy for discretion in financial statements is an estimate of discretionary accruals. In many studies, the accounting quality measure is unsigned (e.g., Francis et al. 2004, 2005; Hribar and Nichols 2007). In other words, "extreme" discretionary accruals of either sign are proxies for accounting numbers that are likely manipulated. A counter-argument is that it is only the extreme positive discretionary accruals (that is, the income-increasing accruals) that lower accounting quality. We have designed our study to explicitly examine that assumption by separating "extreme" negative and "extreme" positive accruals.

Both proxies for discretion, estimated discretionary accruals and the existence of a restatement, reflect a combination of separate factors that relate to many of the financial reporting process. These include judgments within GAAP, managerial incentives, and the costs and benefits of exercising discretion. Our study does not attempt to assess the differential effects of each of these components separately.

3.2 Effects of unrecorded intangible assets

Financial statements do not recognize many forms of intangible assets, such as research and development expenditures, which are generally fully expensed in the year of incurrence. A substantial literature examines the implications of unrecognized intangible assets for the informativeness of financial statements and finds that the financial statements of firms with material intangible assets have lower value relevance. In a security price context, for example, a number of studies document that research and development expenditures are priced and treated as economic assets (for example, Lev and Sougiannis 1996). These findings suggest that the presence of unrecognized intangible assets will reduce the predictive power of bankruptcy models based on accounting ratios. Intangible assets constitute omitted assets whose exclusion from financial statements can induce measurement error in the accounting variables, such as an understatement of assets and net income (for a growing firm). This understatement can lead to an understatement of profitability and an overstatement of leverage.[2] From this perspective, the alternative hypothesis is that those firms with the greatest research and development intensity will be associated with a lower predictive power with respect to the bankruptcy model.

The null hypothesis with respect to intangible assets is that their presence may not lead to deterioration in predictive power because the value of intangible assets either disappears or is nontransferable as bankruptcy approaches. For example, traditional financial statement analysis (for example, Graham and Dodd 1934) focuses on tangible assets, even to the point of eliminating recognized intangibles such as goodwill.

3.3 Book-to-market ratios

We examine the predictive power of bankruptcy models across various categories of the book-to market ratio. The book-to market ratio has been viewed in various ways by prior research, including as a proxy for intangible assets. Here we also view the book-to-market ratio as a partial manifestation of the comprehensiveness of accounting standards. In particular, in a setting where the accounting book value of equity and the market value of equity are identical (for example, comprehensive market-value accounting), the book-to market ratio would be one. The book-to-market ratio can depart from one if economic impairments to asset values are unrecorded, in which case the book-to market ratio is above one, or there are unrecognized increases in economic value of tangibles or unrecognized intangible assets, in which case the book-to-market ratio is below one. Our purpose is to determine if there is differential predictive power in those firm-years where the

[2] The effect of expensing intangibles on profitability depends on the growth of the firm. The effect of unrecognized assets unambiguously increases the leverage ratio.

book-to-market ratios differ most from one. The null hypothesis is, therefore, that there are no differences in predictive power when the book-to-market ratio deviates from one, while the alternate hypothesis is that predictive power is lower. Of course, the book-to- market ratio can proxy for a variety of forces and, hence, the findings regarding the book-to market ratio are open to multiple interpretations. However, because we conduct these analyses in conjunction with other measures of financial reporting quality, we believe they offer additional evidence concerning our basic predictions.

3.4 Recognition of losses

Prior research documents a striking increase in the frequency of losses over time (Collins et al. 1997; Bradshaw and Sloan 2002; Hayn 1995; Givoly and Hayn 2000). A number of researchers suggest the increasing frequency of loss recognition over time reflects increasing conservatism (Hayn 1995, Basu 1997, and Givoly and Hayn 2000). A rationale for this is that accounting standards, such as changes in the impairment standards introduced by SFAS 144 (FASB 2001), require more timely recognition of losses over time. These studies also document that losses are less persistent. The lower degree of persistence could lower the predictive power for loss firms. Of course, the frequency of losses is the joint effect of accounting standards and underlying economic conditions. For example, the economy and certain sectors, such as high tech, may vary in riskiness over time. We do not attempt to disentangle these joint forces. Instead, we examine whether the predictive power of bankruptcy models varies cross-sectionally with the recognition of losses.

Our null hypothesis is that the financial statements of firms recognizing losses do not differ in predictive ability for bankruptcy relative to those of firms not recognizing losses. Our alternate hypothesis is that firms recognizing losses have differential predictive ability, but we do not specify whether loss recognition results in enhanced or impaired predictive ability. One could argue that more timely recognition of losses improves the predictive ability of financial statements for bankruptcy. However, to the extent that loss recognition is discretionary, as with, say, “big baths,” and reflects the ability to take an “earnings hit,” predictive power could be adversely affected by loss recognition. In addition, prior research documents that investors assign different values to the earnings of loss versus profit firms because losses are less persistent than profits. For both these reasons, loss firms could have less informative financial ratios. Our test of this hypothesis is therefore two-tailed.

3.5 Analysis of the accounting, market and combined models

As mentioned earlier, our study examines accounting-based, market-based, and combined models so we can compare predictive ability of each class of explanatory variables. A bankruptcy prediction model based on accounting ratios is subject to measurement error in the explanatory variables. We would therefore expect reduced ability to predict bankruptcy when financial ratios are based on less informative financial statements. In contrast, a bankruptcy prediction model based on market-based variables is not necessarily impaired for firms with less informative financial statements. A key factor is how the financial reporting attribute affects the total mix of information embedded in security prices. Relatedly, it is an open question whether the combined model, drawing on information from financial ratios and market-related variables, is impaired if the accounting model has lower predictive ability. Our tests of differences in predictive ability for the market model and combined model are therefore two-tailed.

4 Description of the estimation model

Following Shumway (2001), we use hazard analysis, also known as survival or duration analysis, as our statistical estimation method. Our sample includes nonbankrupt and bankrupt firms, with the nonbankrupt firms coded zero every year they are in the sample and the bankrupt firms coded zero in every sample year except the year of bankruptcy. As Shumway (2001) notes, an advantage of this approach is improved efficiency and reduced bias in the estimated coefficients relative to a static model with a single firm-year observation for failed and nonfailed firms.

The general form of the hazard model we estimate is as follows:

$$\ln h_j(t) = \alpha(t) + \underline{\mathrm{B}}X_j(t). \tag{1}$$

In this model, h_j (t) represents the hazard, or instantaneous risk of bankruptcy, at time t for company j, conditional on survival to t; α (t) is the baseline hazard; $\underline{\mathrm{B}}$ is a vector of coefficients; X_j (t) is a matrix of observations on financial ratios, market-based variables, or both types of variables, which vary with time. The hazard ratio is defined as the likelihood odds ratio in favor of bankruptcy, and the baseline hazard rate is assumed to be a constant. The model is estimated as a discrete time logit model, using maximum likelihood methods, and provides consistent estimates of the coefficients $\underline{\mathrm{B}}$.

The accounting-based estimation model used in Beaver et al. (2005) includes three accounting based variables, which are return on assets (*ROA*), EBITDA divided by total liabilities (*ETL*), and leverage (*LTA*). Prior research has indicated that the relation between security returns and earnings is nonlinear. In the spirit of Collins et al. (1999), we include an indicator variable, *NROAI*, which is one if *ROA* is negative and 0 otherwise. The indicator variable permits different intercepts and different slopes for loss versus nonloss firm-years.[3]

Market-based variables include a proxy for size (*LRSIZE*), the lagged cumulative security residual return (*LERET*), and the lagged standard deviation of security returns (*LSIGMA*). The combined estimation model includes both accounting-based

[3] Including a measure of persistent losses might improve our ability to predict bankruptcy. Easton et al. (2009) find that persistent losses have a larger association with bond returns than transitory losses. We re-estimated our models including a lagged loss indicator in addition to the loss indicator. We find that this variable is not statistically significant and that its inclusion does not improve the predictive power of the model.

and market-based variables. The construction of these variables is discussed in the next section.

We choose to use a reduced-form model as opposed to a structural model based on Merton (1974) in our main analyses. Untabulated results show that our reduced form model has higher predictive power than a similar model that replaces the market variables by a distance to default measure. These findings are consistent with Campbell et al. (2008) and Bharath and Shumway (2008). Moreover, as discussed in Sect. 6.7, we find that all of our results are robust to this alternative model specification that includes a distance to default measure based on the Merton model.

5 Sample properties and descriptive statistics

5.1 Sample properties

Our sample includes bankrupt and nonbankrupt firms listed on NYSE/AMEX or NASDAQ from 1962 through 2002 (the first years of the sample only include NYSE/AMEX firms, as the NASDAQ subsample only starts in 1973). We combine the bankruptcy database from Beaver et al. (2005), which was derived from multiple sources, including CRSP, Compustat, Bankruptcy.com, *Capital Changes Reporter*, and a list provided by Shumway, with a list of bankrupt firms provided by Chava and Jarrow.[4] By including NASDAQ firms in the sample, our aim is to increase statistical power through a larger sample and greater cross-sectional variation in the explanatory variables.[5] As in prior research, financial and utility firms are excluded from the sample.

All independent variables are lagged to ensure that the data were observable prior to the declaration of bankruptcy. We assume that financial statements are available by the end of the third month after the firm's fiscal year-end. As a result, financial statements for the most recent fiscal year are not assumed to be available for firms declaring bankruptcy within 3 months of their fiscal year-end. In this case, and to ensure that accounting information is observable before bankruptcy is declared, we use accounting data for the preceding fiscal year. This handicaps the accounting model relative to the market model, which includes return and price information for the year prior to bankruptcy.

Table 1 reports that the number of bankrupt firms used in the estimation models is 1,251, of which 487 are listed on NYSE-AMEX and 749 are listed on NASDAQ. The inclusion of NASDAQ firms almost triples the number of bankrupt firms. In addition, the conditional probability of failure for NASDAQ firms (749/69,924) is 1.4 times greater than that of NYSE/AMEX firms (487/64,189).

For each of these observations, we require that the company's PERMNO and the bankruptcy date are available. The CRSP PERMNOs from this sample are then

[4] A description of these samples is in Shumway (2001) and Chava and Jarrow (2004). We greatly appreciate the generosity of Tyler Shumway, Sudheer Chava and Robert Jarrow in making their samples available to us.

[5] 69,845 observations were used in the regression analysis in Beaver et al. (2005). By including NASDAQ firms, we increase sample size to 135,455 observations.

Table 1 Sample selection

	Number of firms			Number of firm years		
	Bankrupt	Nonbankrupt	Total	Bankrupt	Nonbankrupt	Total
Firms with available PERMNO and GVKEY	1,857					
Firms not listed in NYSE, AMEX or NASDAQ	(54)					
Firms listed in NYSE, AMEX or NASDAQ with available identifiers	1,803					
Financial and utilities firms	(179)					
Number of firms	1,624					
Firms with no asset information	(12)					
Firms with available asset information	1,612					
Firms for which accounting model data is unavailable	(211)					
Firm for which accounting variables are available (after filling in)	1,401					
Firms for which market model data is unavailable	(14)					
Number of bankrupt firms used in the analysis	1,387					
Firms for which accounting or market data is unavailable in the bankruptcy year, but available for at least one of the years prior to bankruptcy	(136)					
Pooled sample						
Year of bankruptcy			1,251	1,251		
Earlier bankrupt years			1,387	9,989		
Total	1,387	12,978	14,365	11,240	124,215	135,455
Transition years	164	1,178	1,342			
NYSE/AMEX						
Year of bankruptcy	487			487		
Earlier bankrupt years	560			5,775		
Total	560	4,297	4,857	6,262	57,927	64,189
NASDAQ						
Year of bankruptcy	749	749				
Earlier bankrupt years	972	4,065				
Total	972	9,706	10,678	4,814	65,110	69,924

Our final sample includes bankrupt and nonbankrupt firms listed on NYSE/AMEX or NASDAQ from 1962 through 2002. (The first years of the sample include only NYSE/AMEX firms, as the NASDAQ subsample only starts in 1973.) It is the result of the combination of a sample of bankrupt firms generously provided by Chava and Jarrow and the sample of bankrupt firms in Beaver et al. (2005). There were 2,014 observations referring to 1967 firms in the Chava Jarrow sample. We could match 1,780 of these firms to Compustat and CRSP. Five hundred and four of these matched firms were also included in the Beaver, McNichols, and Rhie (2005) sample (cf. their Table 1), while 1,276 firms were not. Also 77 firms in their sample were not in the Chava Jarrow sample. As a result, we obtained a sample of 1,857 firms with available PERMNO and GVKEY identifiers. The process by which we built the final sample used in the regression analysis, departing from these 1,857 bankrupt firms. Firm-year observations are split between the NYSE/AMEX and NASDAQ subsamples, with transition years being omitted from both samples. Note that the sum of the number of firms in the two subsamples (presented in the first three columns) is higher than the number of firms in the pooled sample. This occurs because firms that have transitioned from one stock exchange to the other will appear in both subsamples matched to those in the Compustat Link History File (crsp.cstlink) and the corresponding Compustat identifiers (GVKEYs) are retrieved. In this process, we obtain a sample of bankrupt firms with available PERMNO and GVKEY information. Moreover, as shown in Table 1, we obtain "nonbankrupt" firms with available PERMNO and GVKEY data through the Compustat Link History File. All firms that did not file for bankruptcy in the sample period are included in the sample as nonbankrupt firms. We require that, in each year, firms are listed in NYSE, AMEX, or NASDAQ and that the CRSP variable EXCHCD is either 1, 2, or 3. We exclude financial and utility firms, as the probability of bankruptcy can rest on regulatory decisions as well as financial condition.

Our tests require data on the accounting and market variables used in the regression analysis. As a result, the sample used in the estimation of the model coefficients includes 1,251 bankrupt firm-years and 135,455 total firm-year observations, with 124,215 firm-year observations of nonbankrupt firms as well as 9,989 firm-year observations of bankrupt firms in years other than the year before bankruptcy.

For part of the analysis, this sample is split in two subsamples: NYSE/AMEX and NASDAQ. As discussed above, the addition of the NASDAQ sample almost triples the number of bankrupt firms in the sample. For those firms that transitioned between these stock exchanges during the sample period, the transition year is excluded from both subsamples. For this reason, in Table 1 the sum of the firm-year observations for the firm-years is less than that of the combined sample.

5.2 Definition of variables and descriptive statistics

Our choice of accounting and market-based explanatory variables is motivated by prior research by Altman (1968); Ohlson (1980); Shumway (2001); Hillegeist et al. (2004); Beaver et al. (2005); and Campbell et al. (2008), among others. We include *ROA* to capture profitability, *ETL* to capture the ability of cash flow from operations pre-interest and pre-tax to cover principal and interest payments, and *LTA* to capture leverage. *ROA* is the return on total assets, defined as earnings before interest adjusted for interest income tax (Compustat data172+data15*(1-tax rate))/lagged data6).[6] *ETL* is net income before interest, taxes, depreciation, depletion and amortization divided by total liabilities, both short term and long term (Compustat data13/data181). *LTA* is the ratio between total liabilities and total assets (Compustat data181/data6). In addition to these variables, we include an indicator variable for negative *ROA* (*NROAI*).

The explanatory variables for the market model include proxies for size (*LRSIZE*), stock market performance (*LERET)*, and volatility (*LSIGMA*). *LRSIZE* is the logarithm of the market capitalization as of the end of the third month after the end of the fiscal year, divided by the market capitalization of the market index of NYSE, AMEX, and NASDAQ firms. *LERET* is the prior year's security returns, where security returns are calculated over a 12-month period ending with the third

[6] We assume there is no tax benefit associated with interest for loss firms. For firms that are profitable, the tax benefit for a given year is calculated based on the maximum statutory tax rate for that year.

month after the end of the fiscal year. *LSIGMA* is the standard deviation of the residual return from a regression of the security's monthly return on the return of the market portfolio (the return for a 12-month period ending with the third month of the fiscal year is used in this regression, to ensure that financial statement information is available). These three market variables are computed based on CRSP data. These variables are more precisely defined in the "Appendix".

Our tests require proxies for four financial reporting attributes: discretionary behavior, the magnitude of unrecognized intangible assets, the comprehensiveness of financial reporting, and the incurrence of losses. Two proxies for discretion are used: the occurrence of restated financial statements in a given firm-year and the magnitude of discretionary accruals. The restatement variable (*DREST*), is equal to one for a given fiscal year if this is a manipulation year and zero otherwise. Restatement years are identified based on the five databases described in the "Appendix". These include two restatement databases (the GAO and Huron databases), two databases containing Accounting and Auditing Enforcement Releases (the database from Bonner et al. 1998, which was generously made available by the authors, and a sample of AAERs hand collected from the SEC website), and one database of class action security lawsuits provided by Woodruff-Sawyer.[7] By combining these five databases, we can obtain the most comprehensive restatement database we are aware of, in terms of number of years covered.

To estimate discretionary accruals, as in Dechow et al. (1995), among others, we run a cross sectional regression of current accruals on change in sales, adjusted by the change in receivables (with the independent and dependent variable scaled by lagged total assets).[8] Through this process we obtain a set of coefficients for each industry and sample year, which we use to estimate nondiscretionary accruals. Discretionary accruals (*DACC*) are then calculated as the difference between total current accruals and nondiscretionary accruals.[9]

As a proxy for unrecognized intangible assets, we compute R&D expenses as a percentage of sales (*RDSALES*, that is, Compustat data46/data12). We then calculate the mean of this measure for each firm, over all years leading up to and including the year in which accounting ratios are measured.[10] Firms are ranked in terms of R&D intensity based on this mean.

Firms are also partitioned based on the book-to-market ratio (*BTM*), which is calculated as the ratio of book value of equity (Compustat data 216) to market capitalization at fiscal year end (Compustat data25*data199). *BTM* is measured in the same period as *ROA* and the other accounting variables. In contrast to most studies, we do not exclude firms with negative book value of equity. As a result,

[7] Woodruff-Sawyer is a full-service insurance brokerage and consulting firm based in San Francisco.

[8] Given that our sample period begins in 1962 and therefore that cash flow statement information is not available for most of the sample, we compute current accruals using a balance sheet approach. In particular, current accruals are equal to the change in current assets minus change in current liabilities and in cash plus the change in short term debt (i.e. Compustat Δdata4-Δdata5-Δdata1+Δdata34).

[9] Our results are robust to an alternative specification of the accrual model, which includes a proxy for the change in cash flows, following Kasznik (1999).

[10] We also repeated the analysis using data for the entire time series available for the firm. The results were essentially the same.

some of the firms in the sample have a negative *BTM* ratio. We compare the predictive power of our models across four main groups of observations: firm years with negative *BTM*, in the top decile of positive *BTM*, in the bottom decile of positive *BTM*, and firms with "medium" *BTM* (that is, neither in the top nor bottom decile).

As discussed earlier, the incidence of losses has been viewed as a proxy for conservatism in financial statements. However, the incurrence of losses is also affected by underlying economic conditions. We measure the incurrence of losses as an indicator variable for negative *ROA* (*NROAI*) and define loss years as years for which *ROA* is negative.

When an accounting variable is missing for a given year, we use its lagged value. We fill in missing values of *DACC*, *RDSALES*, and *BTM* ratios in the same fashion. Variables are winsorized at the 1 and 99 % levels.

Table 2 presents descriptive statistics for the combined sample and for the NYSE/AMEX and NASDAQ samples. There are several striking differences between the NYSE/AMEX and NASDAQ samples. The NASDAQ sample has a higher frequency of losses, as evidenced by the mean of *NROAI* of 35.1 versus 11.5 % for the NYSE/Amex sample. The NASDAQ sample has lower return on assets: −4.2 versus 6.3 % for NYSE/AMEX. Similarly, *EBITDA* to total liabilities, *ETL*, has a mean of −1 versus 32.8 % for NYSE/AMEX. The NASDAQ sample exhibits higher residual return volatility (16 vs. 10 %), smaller market capitalization (−11.81 vs. −9.72 %), a higher frequency of restatements, (1.9 vs. 1.0 %), a higher standard deviation of discretionary accruals, (0.138 vs. 0.091), higher R&D expenditures (13.6 vs. 1.8 %), lower book-to-market ratio (0.81 vs. 0.93), and lower leverage (48.71 vs. 52.88 %).

Even though leverage is lower for the NASDAQ sample, the relative frequency of bankruptcy is higher. For virtually all of the other measures, including the volatility of residual security returns, the risk also would appear to be higher. This difference is consistent with the NASDAQ sample having higher business risk, due to differences in sector composition between NASDAQ and NYSE/AMEX, including the greater frequency of high tech firms in the NASDAQ sample.

Figure 1 presents plots of the explanatory variables over time. *MDREST* is the percentage of sample firms that restated their financial statements for year t; *MHIGHDACC* is the percentage of firms in year t whose discretionary accruals represent more than 10 % of lagged assets; *MHIGHRD* is the percentage of firms whose R&D sales represent more than 5 % of sales in year t; *MNROAI* is the percentage of loss firms; *MBTM* is the percentage of firms in year t in deciles 5 through 8 of book-to-market (where deciles are calculated for the entire sample). The precise definition of the variables is described in the "Appendix". Panel A shows the frequency of restatements from 1962 through 2002 and shows a striking increase through the 1990s and early 2000s. Panel B shows the frequency of high discretionary accruals over time and exhibits an increasing pattern, particularly for NASDAQ firms and the sample as a whole, over time. Panel C shows the frequency of firm-years with high R&D expenses relative to sales and shows a significant increase over time, especially for the NASDAQ sample. Panel D shows the frequency of losses and shows an increasing trend over time for both exchanges and

Table 2 Descriptive statistics

	N	Mean	Median	SD	Minimum	Maximum
Panel A: combined sample						
NROAI	135,455	0.2375	0	0.4256	0	1
ROA	135,455	0.0084	0.0617	0.2337	−1.6503	0.4254
LTA	135,455	0.5073	0.4971	0.2690	0.0397	2.9323
ETL	135,455	0.1526	0.2314	0.9165	−5.9776	2.2235
LERET	135,455	0.0273	0.0089	0.5362	−1.4327	1.8589
LSIGMA	135,455	0.1330	0.1093	0.0896	0.0122	0.5303
SIZE	135,455	490,905	48,686	1,522,923	569	10,769,146
LRSIZE	135,455	−10.8128	−10.9133	2.1235	−15.4133	−5.6305
DREST	135,455	0.0151	0	0.1220	0	1
DACC	130,828	−0.0015	−0.0046	0.1183	−0.4150	0.4423
RDSALES	134,674	0.0792	0	0.3610	0	3.0564
BTM	131,560	0.8636	0.6402	0.8331	−0.6892	4.6518
Panel B: NYSE/AMEX sample						
NROAI	64,189	0.1154	0	0.3195	0	1
ROA	64,189	0.0629	0.0702	0.1118	−1.6171	0.4254
LTA	64,189	0.5288	0.5210	0.2292	0.0397	2.9323
ETL	64,189	0.3281	0.2631	0.4535	−5.9776	2.2235
LERET	64,189	0.0374	0.0224	0.4115	−1.4327	1.8589
LSIGMA	64,189	0.1015	0.0861	0.0625	0.0122	0.5303
SIZE	64,189	817,584	104,516	1,992,950	569	10,769,146
LRSIZE	64,189	−9.7218	−9.7666	1.9766	−15.4133	−5.6305
DREST	64,189	0.0103	0	0.1010	0	1
DACC	61,453	−0.0006	−0.0047	0.0911	−0.4150	0.4423
RDSALES	64,136	0.0185	0	0.1295	0	3.0564
BTM	60,653	0.9283	0.7241	0.8030	−0.6892	4.6518
Panel C: NASDAQ sample						
NROAI	69,924	0.3512	0	0.4774	0	1
ROA	69,924	−0.0424	0.0468	0.2965	−1.6503	0.4254
LTA	69,924	0.4871	0.4637	0.2996	0.0397	2.9323
ETL	69,924	−0.0103	0.1853	1.1696	−5.9776	2.2235
LERET	69,924	0.0177	−0.0103	0.6290	−1.4327	1.8589
LSIGMA	69,924	0.1618	0.1376	0.1006	0.0122	0.5303
SIZE	69,924	193,483	26,833	800,808	569	10,769,146
LRSIZE	69,924	−11.8156	−11.8415	1.7319	−15.4133	−5.6305
DREST	69,924	0.0194	0	0.1381	0	1
DACC	68,069	−0.0028	−0.0047	0.1380	−0.4150	0.4423
RDSALES	69,199	0.1358	0	0.4791	0	3.0564
BTM	69,576	0.8112	0.5657	0.8573	−0.6892	4.6518

Descriptive statistics for the independent variables in the bankruptcy prediction model as well as for the partition variables used in the analysis. All variables are defined in the "Appendix" and winsorized at 1 and 99 %

the sample overall. NASDAQ firms exhibit a significantly higher frequency of losses, beginning in the early 1980s, climbing to over 50 % of firm-years by 2002. Panel E shows the frequency of firm-years with book-to-market ratios in deciles 5 through 8, corresponding to ratios closer to one. As Panel E shows, there is a generally declining tendency for firms to have book-to-market ratios close to one,

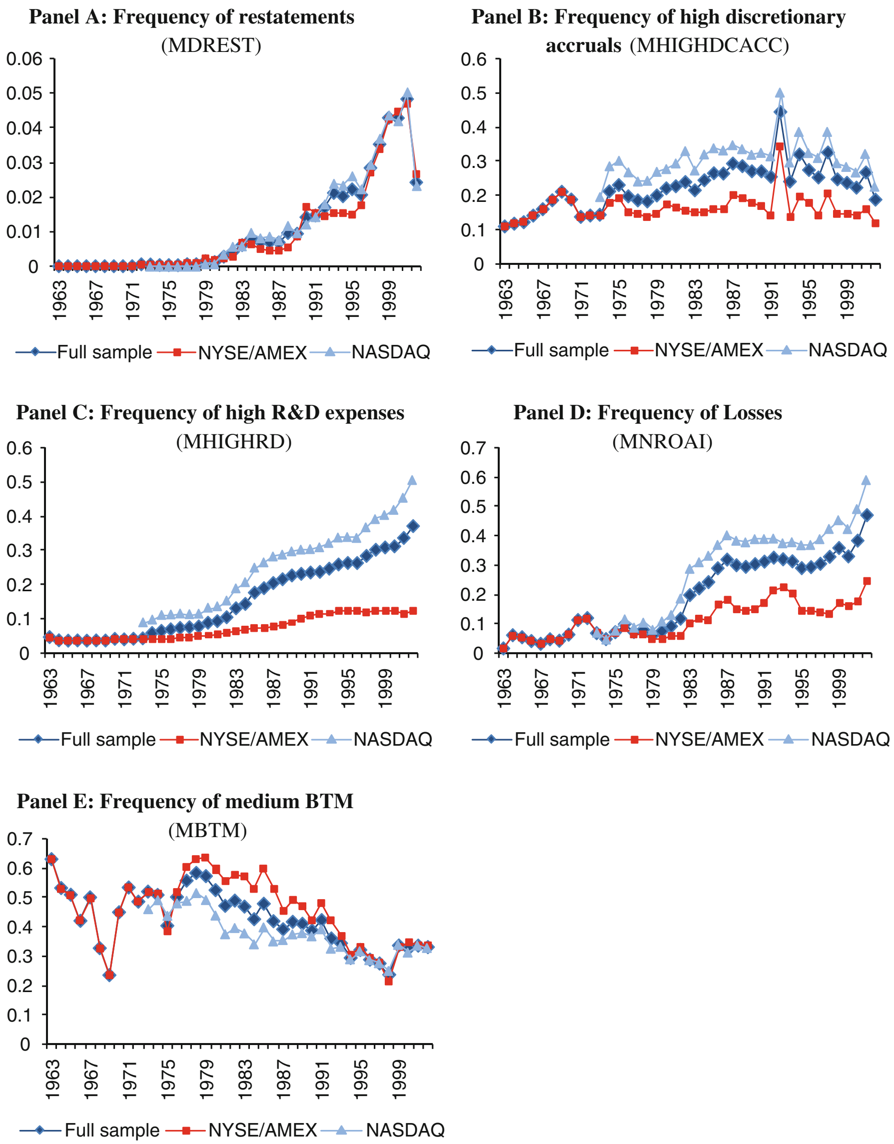

Fig. 1 Plot of explanatory variables over time. This shows the frequency of restatements, high discretionary accruals, high R&D expenses, losses, and medium BTM over time. All variables are defined in the "Appendix"

with firms from both exchanges exhibiting a general decline, and with NASDAQ firms showing a lower frequency of book-to-market ratios close to one from the 1970s to the mid-1990s.

6 Results

6.1 Bankruptcy prediction models

Panel A of Table 3 reports the estimation results for the accounting model. The model includes an indicator variable (*NROAI*) for lack of profitability and separate slope coefficients for the loss firm-years. The coefficient on this indicator variable is significantly positive, which implies the probability of bankruptcy is significantly higher for loss firms. A coefficient of 2.296 implies that the probability of bankruptcy for a firm with losses is approximately 10 times as great as when net income is positive, conditional upon the other variables in the model. The three remaining accounting variables in the prediction model are significant and of the predicted sign for profitable firms.[11] Probability of default is decreasing in profitability, increasing in leverage and decreasing in *EBITDA* relative to total liabilities. However, the incremental coefficients for the loss firms are of the opposite sign and are significant, implying the combined coefficients for the loss firms are driven toward zero. The coefficient on *ROA* is not significantly different from zero for loss firms. The findings indicate that the presence of a loss is a dominating variable and that, conditional on a loss, the magnitude of the loss does not provide additional predictive power.

The predicted scores are ranked and divided into deciles, based on the combined distribution of bankrupt and nonbankrupt firm-years. The percentage of firm-years in these deciles is then reported separately for the year prior to bankruptcy, prior years for bankrupt firms, and firm-years for nonbankrupt firms. Decile 0 has the highest probability of bankruptcy. The percentage of bankrupt firm-years in the three highest bankruptcy risk deciles (that is, deciles 0 through 2) is adopted as a convenient way of comparing predictive ability across models and samples.

We find that 80.02 % of bankrupt firms appear in the three lowest deciles, that is, in the deciles with the highest estimated probability of bankruptcy, compared with an expected 30 % based on the null hypothesis of no predictive power. In the years before bankruptcy, the 42.71 % of firms in the first three deciles is also higher than expected. The fact that the percentage of earlier firm-years of ultimately bankrupt firms is also asymmetrically distributed in the highest risk deciles indicates that these firms had a higher probability of bankruptcy in these earlier years and that the accounting model could partially identify them even several years before bankruptcy.

[11] We also estimate the basic accounting model developed by Beaver, McNichols, and Rhie (2005), which does not include the loss indicator and does not allow for different slopes for loss firms. The results are similar to those reported in the paper, even though the sample has expanded considerably. Consistent with prior results, all three accounting ratios are significant and have the predicted sign.

Table 3 Hazard model estimation

	Coefficients	Chi-Square	*p* value
Panel A: accounting loss model			
Intercept	−6.61	3,220.79	<0.0001
NROAI	2.30	310.02	<0.0001
ROA	−5.75	30.52	<0.0001
LTA	2.78	426.98	<0.0001
ETL	−0.89	102.43	<0.0001
*NROAI*ROA*	5.90	31.66	<0.0001
*NROAI*LTA*	−1.50	101.06	<0.0001
*NROAI*ETL*	1.13	121.05	<0.0001

	Year of bankruptcy		Earlier bankrupt years		Nonbankrupt firms	
	N	Cumulative (%)	N	Cumulative (%)	N	Cumulative (%)
0	708	56.59	1,766	17.68	11,052	8.90
1	195	72.18	1,355	31.24	11,997	18.56
2	98	80.02	1,145	42.71	12,307	28.46
3	112	88.97	1,283	55.55	12,153	38.25
4	51	93.05	972	65.28	12,525	48.33
5	29	95.36	924	74.53	12,597	58.47
6	26	97.44	763	82.17	12,765	68.75
7	13	98.48	692	89.10	12,839	79.08
8	10	99.28	621	95.31	12,922	89.49
9	9	100.00	468	100.00	13,058	100.00
Total	1,251		9,989		124,215	

	Coefficients	Chi-Square	*p* value
Panel B: market model			
Intercept	−7.36	536.85	<0.0001
LERET	−1.48	513.13	<0.0001
LSIGMA	5.45	246.65	<0.0001
LRSIZE	−0.11	17.92	<0.0001

	Year of bankruptcy		Earlier bankrupt years		Nonbankrupt firms	
	N	Cumulative (%)	N	Cumulative (%)	N	Cumulative (%)
0	631	50.44	1,410	14.12	11,485	9.25
1	239	69.54	1,343	27.56	11,965	18.88
2	157	82.09	1,197	39.54	12,196	28.70
3	85	88.89	1,130	50.86	12,333	38.63
4	50	92.89	964	60.51	12,534	48.72
5	31	95.36	925	69.77	12,594	58.86
6	26	97.44	810	77.88	12,718	69.09
7	9	98.16	737	85.25	12,798	79.40
8	11	99.04	704	92.30	12,838	89.73
9	12	100.00	769	100.00	12,754	100.00
Total	1,251		9,989		124,215	

Table 3 continued

	Coefficients	Chi-Square	*p* value
Panel C: combined loss model			
Intercept	−9.43	629.90	<0.0001
NROAI	4.00	73.17	<0.0001
ROA	−2.61	8.88	0.0029
LTA	2.36	280.81	<0.0001
ETL	−0.56	26.29	<0.0001
LERET	−1.87	304.49	<0.0001
LSIGMA	7.80	153.98	<0.0001
LRSIZE	−0.13	14.48	0.0001
*NROAI*ROA*	3.24	13.33	0.0003
*NROAI*LTA*	−1.17	56.00	<0.0001
*NROAI*ETL*	0.87	49.40	<0.0001
*NROAI*LERET*	0.68	30.49	<0.0001
*NROAI*LSIGMA*	−3.59	25.01	<0.0001
*NROAI*LRSIZE*	0.13	9.40	0.0022

	Year of bankruptcy		Earlier bankrupt years		Nonbankrupt firms	
	N	Cumulative (%)	N	Cumulative (%)	N	Cumulative (%)
0	798	63.79	1,740	17.42	10,988	8.85
1	216	81.06	1,469	32.13	11,862	18.40
2	113	90.09	1,232	44.46	12,205	28.22
3	51	94.16	1,139	55.86	12,358	38.17
4	29	96.48	1,038	66.25	12,481	48.22
5	19	98.00	894	75.20	12,637	58.39
6	7	98.56	731	82.52	12,816	68.71
7	9	99.28	650	89.03	12,885	79.08
8	4	99.60	601	95.04	12,948	89.51
9	5	100.00	495	100.00	13,035	100.00
Total	1,251		9,989		124,215	

The estimation results for different hazard models for the full sample of bankrupt and nonbankrupt firm-years and the distribution of firms across deciles of the predicted probability of bankruptcy, ranked from the highest probability (decile 0) to lowest probability (decile 9). Panel A presents this analysis for the accounting model. Panels B and C contain the analysis for the market-based hazard model and the combined model. All variables are defined in the "Appendix" (N = 135,455)

Panel B presents the estimated coefficients for the market model. Coefficients have the predicted signs, are significant, and are consistent with those reported in prior research. In particular, the probability of bankruptcy is increasing in volatility of residual returns and decreasing in size and lagged residual return. Moreover, 82.1 % of firms are in the bottom 3 deciles in the year of bankruptcy, which is much greater than the 30 % expected under the null hypothesis of no predictive ability. The classification accuracy is slightly greater than that of the accounting model at 80.02 %, so virtually all of the predictive ability of market-based variables is captured by the three accounting-based variables.

Panel C of Table 3 reports the estimation results for a combined hazard model that includes both accounting and market-based variables. The coefficient on *NROAI*, 4.004, is significant and implies that the presence of a loss implies a firm is more than 50 times as likely to declare bankruptcy. Moreover, even in the presence of the market-based variables, the accounting-based variables remain significant for the profitable firms. This is important because the market-based variables reflect the total mix of information of which financial statements are only a subset and, in principle, could subsume the predictive ability of the accounting-based variables. Similar to the results for the accounting-based model, all of the incremental slope coefficients are of the opposite sign, therefore driving the sum of the respective coefficients toward zero. Despite this fact, untabulated findings indicate that all variables are significant for loss firms, with the exception of *LRSIZE*.

The percentage of bankrupt firm-years in the bottom three deciles is 90.09, which is higher than that for either the accounting or market-based model, consistent with both the accounting and market variables having significant explanatory power.[12] In an efficient capital market, the market-based model should dominate the accounting model, since the total mix of information includes financial statements as a subset. However, the results indicate that approximately the same predictive power is captured by the accounting variables. Moreover, accounting variables provide some explanatory power not provided by market variables. The latter could reflect misspecification of the market variables rather than evidence of market inefficiency. Conversely, the market variables capture some information not captured by the accounting variables. This could reflect information aggregated in prices that does not derive from financial statements as well as possible misspecification of the accounting variables.

6.2 Discretionary behavior results

6.2.1 Earnings restatements

To compare the predictive power of the models for restated and nonrestated years, we rank the hazard scores for all observations within each of the two subsamples by year. These hazard scores are computed based on the pooled estimation of each of the models. (The negative ROA indicator variable is included in this estimation for the accounting and combined models.)

Table 4 contains the frequency of firms in each of the lowest three deciles ranked on hazard scores, which correspond to the highest probability of bankruptcy. As hypothesized, the predictive power of the accounting model is lower for the firm-years where restatement is involved, with only 50.45 % of bankrupt firms in the lowest three deciles for the restatement subsample, in contrast to 82.02 % for nonrestated years. Untabulated statistical tests indicate this difference is significant with a probability value less than 0.01.[13] This finding is consistent with the

[12] Similar results are obtained, in untabulated analysis, for the combined model without separate slopes or indicators but with a slightly lower predictive power of 88.57%.

[13] We assess the significance of the difference using a χ^2 test. In all further comparisons, we refer to a subsample as having higher or lower predictive ability if the difference is significant with $p < 0.01$.

Table 4 Hazard deciles for restatement partition

	Year of bankruptcy		Earlier bankrupt years		Nonbankrupt firms	
	N	Cumulative (%)	N	Cumulative (%)	N	Cumulative (%)
Panel A: accounting model						
Restated years						
0	27	24.32	31	12.70	133	7.87
1	17	39.64	29	24.59	159	17.27
2	12	50.45	20	32.79	174	27.56
Other						
0	672	58.95	1,725	17.70	10,927	8.92
1	176	74.39	1,324	31.29	11,842	18.58
2	87	82.02	1,136	42.95	12,122	28.48
Panel B: market model						
Restated years						
0	39	35.14	20	8.20	132	7.81
1	16	49.55	19	15.98	170	17.86
2	16	63.96	24	25.82	166	27.68
Other						
0	580	50.88	1,342	13.77	11,402	9.31
1	223	70.44	1,320	27.32	11,799	18.94
2	148	83.42	1,184	39.47	12,013	28.74
Panel C: combined model						
Restated years						
0	41	36.94	29	11.89	121	7.16
1	17	52.25	23	21.31	165	16.91
2	18	68.47	19	29.10	169	26.91
Other						
0	740	64.91	1,702	17.47	10,882	8.88
1	196	82.11	1,433	32.17	11,713	18.44
2	100	90.88	1,197	44.45	12,048	28.27

The distribution of firms in the top three deciles of predicted probability of bankruptcy for two groups of observations: *restated years* (obtained using the databases described in the "Appendix") and *other*. The predicted probability of bankruptcy is obtained by using the coefficients from the hazard model for the full sample, presented in Table 3, panels A, B, and C, and ranked within each of the subgroups. All variables are defined in the "Appendix" (N = 135,455)

contention that accounting numbers that are departures from GAAP are of lower quality for bankruptcy prediction. Interestingly, the lower predictive power holds for the market-based model as well, though to a lesser degree. For the restatement firm-years, 63.96 % of bankrupt firm years are in the bottom three deciles, while for the nonrestatement years it is 83.42 %. In other words, even though these variables are based upon the total mix of information, of which accounting data is a subset, their predictive power is also lower. The finding indicates that financial ratios based on manipulated financial statements are less informative for predicting bankruptcy. Furthermore, investors can partially, but not completely, compensate for the less informative financial ratios through other information sources.

Moreover, the differences in predictive power are also present in the combined model. The percentage of bankrupt firms in the first three deciles is 68.47 for the restatement subsample in the combined model, in contrast to 90.88 for the subsample without restatements. The findings indicate that our proxy for the exercise of accounting discretion is associated with lower ability to predict bankruptcy using market and accounting information.

6.2.2 *Discretionary accruals*

Table 5 presents the results for our partition based on discretionary accruals and indicates these findings are similar to those based on restatements. The predictive power of the three models is highest for those firms with a medium level of accruals.

Prior literature has raised the question of whether earnings quality is reduced only for those increasing earnings or whether quality is lower for "extreme" accruals that increase or decrease earnings (for example, Francis et al. 2004). Note that firms with significant amounts of impairments and special charges will likely fall into the bottom *DACC* decile. The results in Table 5 suggest the predictive power of the accounting model is reduced for both low (for example, negative) and high accruals. In the accounting model, for example, the predictive power is greatest for the mid-range of accruals, with 81.82 % in the bottom three deciles and lower for both extremes. The deterioration appears greatest in the highest accrual decile (68.33 % in the bottom three deciles) as compared with the lowest accrual decile (75.93 % in the bottom three deciles), though this difference is only marginally significant (with probability value 0.12).

One could argue that the market-based models would be insensitive to the partitioning on accounting, since they are not directly affected by the magnitude of the accruals. Moreover, the total mix of information may compensate for the effects of extreme accruals because it includes a potentially richer set of information. Unlike restatements, for example, the discretionary portion of accruals is estimable using past and current data. However, the pattern observed in Panel B for the market-based variables is similar and, if anything, more dramatic than for the accounting based model. The middle group has the highest predictive power (86.23 %) while the two extreme accrual deciles have the lowest (65.43 and 65.0 %, respectively). Hence, the information environment for the extreme accrual firm-years appears to be considerably different from that of the mid-range accrual firm-years, and the differences are even more striking for the market-based variables than for accounting-based variables. Extreme accruals may proxy for some other underlying economic difference, not explicitly captured by the market-based variables, that makes bankruptcy prediction more difficult. Here, the effects for the low and high accrual groups are symmetric.

The results in Panel C of Table 5 for the combined model show the same pattern, less pronounced than for the market model but more pronounced than for the accounting model, with a symmetric pattern for low and high accruals groups, as in

Table 5 Hazard deciles for discretionary accruals partition

	Year of bankruptcy		Earlier bankrupt years		Nonbankrupt firms	
	N	Cumulative (%)	N	Cumulative (%)	N	Cumulative (%)
Panel A: accounting model						
Low DACC						
0	115	35.49	145	12.80	1,033	8.89
1	77	59.26	148	25.86	1,085	18.22
2	54	75.93	138	38.04	1,123	27.88
Medium DACC						
0	424	56.68	1,338	18.82	8,688	8.97
1	107	70.99	967	32.41	9,395	18.68
2	81	81.82	913	45.25	9,476	28.47
High DACC						
0	51	42.50	168	13.18	1,069	9.15
1	17	56.67	146	24.63	1,150	18.99
2	14	68.33	151	36.47	1,149	28.82
Panel B: market model						
Low DACC						
0	98	30.25	108	9.53	1,087	9.35
1	69	51.54	122	20.30	1,119	18.98
2	45	65.43	135	32.22	1,135	28.74
Medium DACC						
0	416	55.61	1,063	14.95	8,971	9.27
1	147	75.27	967	28.55	9,355	18.93
2	82	86.23	921	41.50	9,467	28.71
High DACC						
0	35	29.17	140	10.98	1,113	9.52
1	24	49.17	142	22.12	1,147	19.34
2	19	65.00	133	32.55	1,162	29.28
Panel C: combined model						
Low DACC						
0	128	39.51	131	11.56	1,034	8.89
1	72	61.73	149	24.71	1,089	18.26
2	44	75.31	139	36.98	1,132	28.00
Medium DACC						
0	516	68.98	1,351	19.00	8,583	8.87
1	116	84.49	1,045	33.69	9,308	18.48
2	56	91.98	918	46.60	9,496	28.29
High DACC						
0	53	44.17	180	14.12	1,055	9.03
1	19	60.00	144	25.41	1,150	18.87
2	18	75.00	151	37.25	1,145	28.66

The distribution of firms in the top three deciles of predicted probability of bankruptcy for three groups of observations: *Low DACC*, *Medium DACC*, and *High DACC*. The *Low* and *High DACC* categories include firm-years in the bottom and top deciles of discretionary accruals, respectively. *Medium DACC* includes the remaining firm years. The predicted probability of bankruptcy is obtained by using the coefficients from the hazard model for the full sample, presented in Table 3, panels A, B, and C, and ranked within each of the subgroups (N = 130,828). All variables are defined in the "Appendix" the market model. The most extreme positive accruals are of the lowest quality, consistent with overstated earnings being less informative than unbiased or understated earnings. However, consistent with ongoing research on "earnings quality," our findings indicate extreme accruals of either sign are of lower quality with respect to bankruptcy prediction.

6.3 Unrecorded intangible assets

Our measure of intangible assets is based on the ratio of R&D to sales (*R&D/SALES*). Firm-years are partitioned into three groups: firms with zero R&D, firms in the top decile of *R&D/SALES* (high R&D), and all other firms (medium R&D). The hypothesis is that the presence of intangible assets lowers the predictive power of the model because it represents an asset not captured by the accounting variables. Moreover, under the total mix of information, there may be no difference in predictive ability to the extent that the market can price the value of the intangible assets based on the total mix of information.

As reported in Table 6, the predictive power for the accounting model is lowest for the high R&D group at 63.03 %, compared with 89.91 % for the middle group, consistent with the hypothesis. However, the zero R&D group has a somewhat lower predictive power (80.0 %), which is not predicted by the hypothesis. The results for high R&D firms are consistent with Franzen et al. (2007), who find that the O-score is more likely to misclassify a solvent firm with large R&D expenses.

Further, for the market-based model, the lowest predictive power is also in the highest R&D group, though the market model is more informative than the accounting model for these firms. The latter finding is consistent with prices reflecting additional information about the value of intangibles beyond that reflected in financial ratios. The findings are consistent with the interpretation that the market partially prices the intangible asset (as in Barth et al. 1998, Lev 2001, and Lev and Sougiannis 1996). Note that unlike the accounting model, there is only a slight difference for the zero and the middle R&D group. The combined model exhibits essentially the same behavior as the market model.

6.4 Book-to-market results

Table 7 compares the predictive power of the bankruptcy model for negative, low, medium, and high levels of *BTM*. The findings indicate that predictive power differs across categories of the book-to-market ratio. In particular, predictive power is lowest for the firms with a negative *BTM* ratio, that is, negative book value of equity and highest for deciles 1 through 8 of positive *BTM*.

The behavior of the *BTM* ratio is complex. As the probability of bankruptcy increases, both the book value of equity and the market value of common equity decline. It is difficult to predict how the ratio of the two should behave in part because it is difficult to predict which component will decline at a more rapid rate. Moreover, book value can be negative and approaching zero, while market value cannot. In particular, the option value of common equity can remain even as the probability of bankruptcy

Table 6 Hazard deciles for R&D partition

	Year of bankruptcy		Earlier bankrupt years		Nonbankrupt firms	
	N	Cumulative (%)	N	Cumulative (%)	N	Cumulative (%)
Panel A: accounting model						
Average R&D = 0						
0	425	52.80	893	16.48	4,545	8.64
1	140	70.19	690	29.22	5,058	18.25
2	79	80.00	672	41.62	5,139	28.02
Medium R&D						
0	200	61.73	761	18.44	5,295	9.08
1	66	82.10	595	32.86	5,617	18.72
2	25	89.81	563	46.50	5,696	28.49
High R&D						
0	41	34.45	53	11.94	1,244	9.61
1	18	49.58	61	25.68	1,272	19.43
2	16	63.03	60	39.19	1,277	29.29
Panel B: market model						
Average R&D = 0						
0	388	48.20	687	12.68	4,788	9.10
1	165	68.70	725	26.06	4,998	18.60
2	113	82.73	686	38.72	5,091	28.27
Medium R&D						
0	191	58.95	604	14.64	5,461	9.37
1	49	74.07	558	28.16	5,671	19.09
2	41	86.73	513	40.59	5,730	28.92
High R&D						
0	43	36.13	49	11.04	1,246	9.62
1	27	58.82	61	24.77	1,263	19.38
2	17	73.11	50	36.04	1,286	29.31
Panel C: combined model						
Average R&D = 0						
0	479	59.50	861	15.89	4,523	8.60
1	164	79.88	764	29.99	4,960	18.02
2	74	89.07	677	42.49	5,139	27.79
Medium R&D						
0	232	71.60	752	18.22	5,272	9.04
1	46	85.80	618	33.20	5,614	18.67
2	22	92.59	553	46.60	5,709	28.47
High R&D						
0	44	36.97	60	13.51	1,234	9.53
1	27	59.66	61	27.25	1,263	19.28
2	14	71.43	55	39.64	1,284	29.20

The distribution of firms in the top three deciles of predicted probability of bankruptcy for three groups of observations: *Average R&D* = 0, *Medium R&D*, and *High R&D*. The *Average R&D* = 0 category includes observations referring to firms whose average of R&D expenses as a percentage of sales over all the firm-years leading up to and including the year for which financial ratios are calculated. The *High R&D* category includes firms whose average R&D is in the top decile. All remaining observations are included in the *Medium R&D* category. The predicted probability of bankruptcy is obtained by using the coefficients from the hazard model for the full sample, presented in Table 3, panels A, B, and C and ranked within each of the subgroups. All variables are defined in the "Appendix" (N = 134,674)

Table 7 Hazard deciles for BTM partition

	Year of bankruptcy		Earlier bankrupt years		Nonbankrupt firms	
	N	Cumulative (%)	N	Cumulative (%)	N	Cumulative (%)
Panel A: accounting model						
Negative BTM						
0	25	8.56	23	6.18	289	10.06
1	27	17.81	39	16.67	289	20.13
2	36	30.14	48	29.57	272	29.60
Low positive BTM						
0	48	37.21	136	18.40	1,080	9.05
1	26	57.36	100	31.94	1,156	18.74
2	13	67.44	92	44.38	1,181	28.63
Medium positive BTM						
0	313	55.99	1,218	17.19	8,692	9.17
1	58	66.37	860	29.33	9,325	19.01
2	71	79.07	979	43.14	9,199	28.72
High positive BTM						
0	72	31.72	196	14.44	990	8.83
1	33	46.26	188	28.30	1,061	18.28
2	28	58.59	187	42.08	1,072	27.84
Panel B: market model						
Negative BTM						
0	57	19.52	24	6.45	256	8.91
1	68	42.81	23	12.63	264	18.11
2	35	54.79	43	24.19	278	27.79
Low positive BTM						
0	50	38.76	91	12.31	1,123	9.41
1	29	61.24	89	24.36	1,164	19.16
2	15	72.87	85	35.86	1,186	29.10
Medium positive BTM						
0	281	50.27	1,001	14.13	8,941	9.43
1	114	70.66	926	27.19	9,203	19.14
2	62	81.75	861	39.35	9,326	28.98
High positive BTM						
0	67	29.52	160	11.79	1,031	9.19
1	41	47.58	149	22.77	1,092	18.92
2	30	60.79	145	33.46	1,112	28.84

Table 7 continued

	Year of bankruptcy		Earlier bankrupt years		Nonbankrupt firms	
	N	Cumulative (%)	N	Cumulative (%)	N	Cumulative (%)
Panel C: combined model						
Negative BTM						
0	55	18.84	37	9.95	245	8.53
1	59	39.04	32	18.55	264	17.72
2	47	55.14	39	29.03	270	27.12
Low positive BTM						
0	62	48.06	118	15.97	1,084	9.08
1	21	64.34	99	29.36	1,162	18.82
2	12	73.64	108	43.98	1,166	28.59
Medium positive BTM						
0	365	65.30	1,241	17.51	8,617	9.09
1	74	78.53	980	31.34	9,189	18.79
2	53	88.01	881	43.78	9,315	28.62
High positive BTM						
0	86	37.89	187	13.78	985	8.78
1	48	59.03	191	27.86	1,043	18.08
2	33	73.57	177	40.90	1,077	27.68

The distribution of firms in the top three deciles of predicted probability of bankruptcy for four groups of observations: *Negative BTM*, *Low positive BTM*, *Medium positive BTM*, and *High positive BTM*. The *Negative BTM* category includes firm-years with negative BTM. The *Low positive* and *High positive BTM* categories include firm-years in the bottom and top deciles of BTM, respectively. The *Medium positive BTM* category includes all remaining observations. All variables are defined in the "Appendix" (N = 131,560)

rises. As a result, as losses cumulate and as book value heads towards zero, the book-to-market ratio can approach zero (the market-to-book ratio approaches infinity).

To partially address these concerns, we partition the book-to-market ratios into four groups: negative, low positive, medium positive, and high positive book-to-market ratios. Table 7 shows that the same pattern of predictive power is exhibited by all three models. The subsamples with lowest predictive power to highest are the negative *BTM*, high positive *BTM*, low positive *BTM*, and medium positive *BTM*. The negative group includes firm-years with negative book value and for whom the probability of bankruptcy would be expected to be high. These are firms with negative book value but positive market value, in part due to the option-like properties of common stock for limited liability corporations as well as possible unrecognized intangible assets. The high positive group includes firms for which the market is recognizing asset impairments but the accounting is not, or at least not to the same extent. The low positive group is a diverse group of firm-years where either the option value of market price is causing the market-to-book ratio to approach infinity for a low book value firm or the firms have substantial unrecognized intangible assets. All three groups would be expected to have—and in fact do have—higher than average probability of bankruptcy.

Overall, the results indicate that when book-to-market ratios have the greatest departure from one, the predictive power of the bankruptcy models is weakest. In fact, for firms with negative book-to-market ratios, the fraction of bankrupt firms in the lowest three deciles of the hazard score based on the accounting model is 30.14 % or approximately what would be expected by chance, suggesting financial ratios are uninformative for these firms. The market model is also least informative for these firms relative to the other book-to-market samples, though more informative than the accounting model, with 54.79 % of bankrupt firms classified in the bottom three deciles of the hazard score. Relatedly, the predictive power is greater for those firm-years in which accounting and market-based measures of value correspond more closely, with 79.07 % correct classification for the firms with book-to-market ratios closest to one. These findings indicate that investors cannot compensate for impaired financial reporting through other information sources. These differences in predictive ability across book-to-market classes may also reflect underlying economic factors that are not captured by the predictive variables or less informative prices for other reasons.

6.5 Predictive power of models for loss firms

We established that the probability of bankruptcy conditional upon a loss is significantly higher in Sect. 6.1. In this section, we examine whether conditioning for the presence of the loss, the predictive power of the models are the same for loss versus nonloss firms. Table 8 reports the percentage of bankrupt firm-years in the bottom three deciles for loss versus nonloss firm-years. The predictive power of the accounting model for the loss firm years is substantially lower (70.75 %) than for the nonloss firm-years (77.62 %).

This finding is consistent with the results discussed earlier which showed that, conditional upon the presence or absence of a loss, the incremental explanatory power of the remaining accounting and market variables is substantially lower for loss firms. Hence, for these firms, additional variables do not provide much information for distinguishing between the probability of failure among the set of loss firms. For nonloss firm-years, while their conditional probability of bankruptcy is lower, the incremental explanatory power of the additional variables is much greater in distinguishing differences in the probability of bankruptcy.

As we have seen in prior results, the differences observed in the accounting model do not disappear in the market model or combined model. In fact, for the market model, the number classified in the bottom three deciles is 61.3 versus 83.4 % for nonloss years. Interestingly, the market model performs less well than the accounting model for firms with losses, similar to our earlier findings for firms with the lowest discretionary accruals. These findings suggest that market variables do not fully convey the information available in financial ratios about these firms.

Table 8 Hazard deciles for loss partition

	Year of bankruptcy		Earlier bankrupt years		Nonbankrupt firms	
	N	Cumulative (%)	N	Cumulative (%)	N	Cumulative (%)
Panel A: accounting model						
Loss years						
0	297	33.41	357	14.65	2,546	8.83
1	198	55.68	354	29.18	2,668	18.08
2	134	70.75	336	42.96	2,751	27.61
Other						
0	191	52.76	1,516	20.07	8,603	9.02
1	63	70.17	1,111	34.79	9,157	18.62
2	27	77.62	874	46.36	9,433	28.51
Panel B: market model						
Loss years						
0	269	30.26	207	8.49	2,724	9.44
1	150	47.13	254	18.92	2,816	19.21
2	126	61.30	254	29.34	2,841	29.06
Other						
0	207	57.18	1,183	15.66	8,920	9.35
1	64	74.86	1,018	29.14	9,249	19.05
2	31	83.43	936	41.54	9,367	28.87
Panel C: combined model						
Loss years						
0	362	40.72	330	13.54	2,508	8.70
1	186	61.64	330	27.08	2,704	18.07
2	116	74.69	285	38.78	2,820	27.85
Other						
0	254	70.17	1,545	20.46	8,511	8.92
1	46	82.87	1,110	35.16	9,175	18.54
2	25	89.78	932	47.50	9,377	28.38

The distribution of firms in the top three deciles of the predicted probability of bankruptcy for two groups of observations: *loss years* (i.e., years with negative ROA) and *other*. The predicted probability of bankruptcy is obtained by using the coefficients from the hazard model for the full sample, presented in Table 3, panels A, B, and C and ranked within each of the subgroups. All variables are defined in the "Appendix" (N = 135,455)

6.6 Time-series analysis

Having found substantial differences cross-sectionally in the predictive power of bankruptcy models based on the presence of discretionary behavior, unrecognized intangible assets, book-to-market ratios, and the incurrence of losses, we apply the time-series approach of Beaver et al. (2005). They found a slight but statistically insignificant decline in the predictive power of the accounting model, slight improvement in the market model, and essentially no time trend in the combined model.

We regress the percentage of bankrupt firm-years in the bottom three deciles of the hazard score in a given calendar year on time. This regression tests whether the fraction of bankrupt firm-years the accounting model correctly classifies as having the highest probability of bankruptcy varies with time. Table 9 reports the estimation results of our time-series regressions. In contrast to the earlier study, we find a decline in the predictive power of the accounting model, no deterioration in the market model, and an overall decline in the combined model.

To provide a visual perspective on these changes over time, Fig. 2 plots the percentage of bankrupt firm-years classified in the bottom 3 deciles of the hazard score for the accounting, market and combined models, by exchange and for the sample as a whole. Panel A confirms a lower level of predictive ability for NASDAQ firms and a decline in classification accuracy for NASDAQ firms and the sample as a whole. Panel B, by contrast, documents relatively consistent classification accuracy for the market model. Although the market model has lower classification accuracy for NASDAQ firms than NYSE/AMEX firms, the accuracy for both exchanges and the sample as a whole is fairly constant over time. Panel C shows the classification accuracy for the combined model over time and suggests erosion in accuracy over time for both the NASDAQ and NYSE/AMEX samples.

Similar to industry and size, time is a generic proxy variable that often fails to provide insight into the underlying factors. However, based on our cross-sectional analysis, we are now in a position to specify the variables for which time may be a proxy. We conduct a regression of the percentage of bankrupt firm-years classified in the bottom 3 deciles of the hazard score on percentage of restatements, the percentage of firms for which the absolute value of discretionary accruals exceeds 10 % of lagged assets, the percentage of high intensity R&D firms (proxied by R&D greater than 5 % of sales), the frequency of book-to-market values close to one, and the percentage of loss firms in a given calendar year. As reported in Panel A of Table 9, all of the explanatory variables are highly correlated making individual contributions difficult to assess. However, Panel B shows the accounting model's lower predictive power occurs in years when there is a larger frequency of restatements, a relatively large amount of discretionary accruals, relatively high research and development intensity, a higher frequency of firms with book-to-market ratios further from one, and a higher frequency of losses. This evidence is consistent with the cross-sectional analysis and helps to identify at least some of the factors associated with the observed decline over time in predictive power. Interestingly, as Panel C shows, the market model exhibits no such decline over time. With the exception of restatements, the predictive ability of the market model is not affected by the variation of accounting quality over time. However, as Panel D shows, the differential predictive ability of the combined model declines significantly over time. This finding suggests that the erosion in predictive ability of financial ratios was not offset by information reflected in the market-related variables.

Table 9 Time series regressions

	T	*MDREST*	*MHIGHDCACC*	*MHIGHRD*	*MBTM*	*MNROAI*
Panel A: Spearman and Pearson correlation matrix						
T	1.0000	0.8726	0.6552	0.9741	−0.8393	0.9182
		<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
MDREST	0.9842	1.0000	0.4256	0.8907	−0.8162	0.8009
	<0.0001		0.0108	<0.0001	<0.0001	<0.0001
MHIGHDCACC	0.6846	0.6804	1	0.6118	−0.6057	0.6142
	<0.0001	<0.0001		<0.0001	0.0001	<0.0001
MHIGHRD	0.9992	0.9847	0.6880	1.0000	−0.8711	0.9646
	<0.0001	<0.0001	<0.0001		<0.0001	<0.0001
MBTM	−0.8532	−0.8369	−0.6980	−0.8552	1.0000	−0.8320
	<0.0001	<0.0001	<0.0001	<0.0001		<0.0001
MNROAI	0.9140	0.8990	0.6356	0.9120	−0.8132	1.0000
	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	

	(1)	(2)	(3)	(4)	(5)	(6)
Panel B: accounting model						
Intercept	1.0913	0.9104	1.0257	1.0123	0.4578	1.0008
	(48.76)	(53.21)	(11.28)	(48.37)	(7.34)	(49.79)
MDREST		−4.7567				
		(−6.1)				
MHIGHDCACC			−0.7252			
			(−1.95)			
MHIGHRD				−0.8869		
				(−10.69)		
MBTM					0.8998	
					(6.18)	
MNROAI						−0.6610
						(−8.63)
T	−0.0102					
	(−13.01)					
R^2	0.7757	0.4762	0.0897	0.7308	0.6216	0.6227
N	28	31	29	28	29	30

Table 9 continued

	(1)	(2)	(3)	(4)	(5)	(6)
Panel C: market model						
Intercept	0.8458	0.8423	0.8270	0.8411	0.8276	0.8400
	(30.34)	(58.85)	(23.12)	(38.03)	(21.16)	(33.39)
MDREST		−1.3189				
		(−2.79)				
MHIGHDCACC			0.0041			
			(0.03)			
MHIGHRD				−0.0724		
				(0.76)		
MBTM					−0.0082	
					(−0.09)	
MNROAI						−0.0507
						(−0.57)
T	−0.0007					
	(−0.74)					
R^2	−0.0171	0.0447	−0.0357	−0.0133	−0.0368	−0.0225
N	30	29	30	30	29	30
Panel D: combined model						
Intercept	1.0203	0.9542	1.0952	0.9893	0.7507	0.9911
	(63.25)	(90.05)	(25.98)	(73.00)	(21.54)	(69.34)
MDREST		−2.5500				
		(−4.42)				
MHIGHDCACC			−0.7326			
			(−4.06)			
MHIGHRD				−0.3793		
				(−5.98)		
MBTM					0.3973	
					(4.78)	
MNROAI						−0.3136
						(−5.7)
T	−0.0042					
	(−6.52)					
R^2	0.4782	0.3858	0.3253	0.4531	0.386	0.4474
N	31	29	30	31	30	31

This table presents the results from the regression of the number of bankruptcies in the top three deciles of the predicted hazard rate on the annual percentage restatements, high discretionary accruals, high R&D, medium BTM and loss firms, and a time trend. Only years with more than five bankruptcies were included in the analysis. Criteria for outlier deletion were as follows: absolute value of studentized residuals greater or equal to 2, Cook's distance greater than the 50 % point of the F distribution with $p + 1$ and $(n - p - 1)$ degrees of freedom, where p is the number of regressors, excluding the constant term and n is the number of observations, absolute value of DFFITS larger than $2\sqrt{(p+1)/(n-p-1)}$. All variables are defined in the "Appendix"

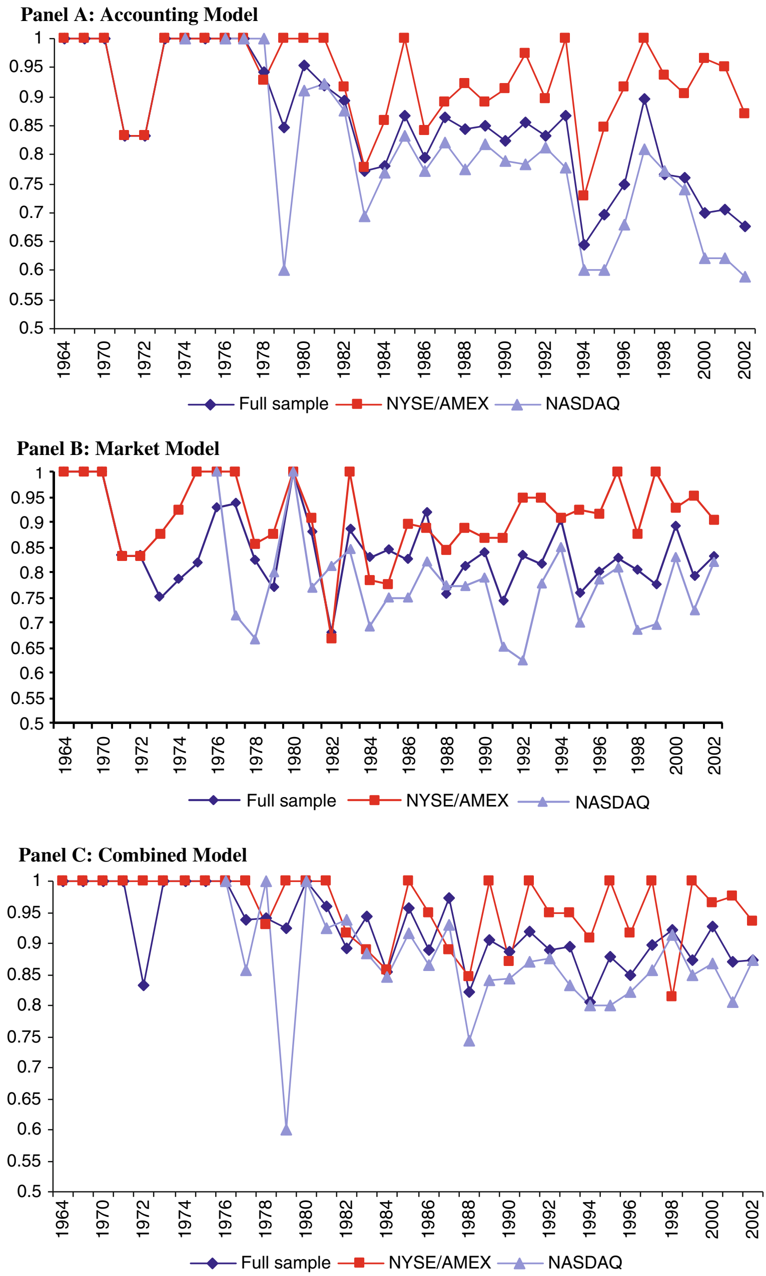

Fig. 2 Percentage of bankrupt firms in the top 3 deciles of predicted hazard rate over time. This shows the percentage of bankrupt firms in the top 3 deciles of the predicted hazard rate over time. Hazard rates are estimated based on the models in **a**, **b**, and **c** of Table 3

6.7 Sensitivity analysis

There is a long tradition in the bankruptcy prediction literature of using out-of-sample testing to avoid a bias of ex post overfitting the data. Shumway (2001) and Beaver et al. (2005), for example, adopt out-of-sample testing. However, such a research design is by no means the predominant research design. Neither of the two more recent contributions—Franzen. Rodgers, and Simin (2007) and Campbell et al. (2008)— employ out-of-sample testing. In any event, the results of out-of-sample testing are reported in Table 10. We randomly divide the overall sample into two subsamples, A and B. We then report four out-of-sample tests as described in detail in Table 10. Of course, there is some slight deterioration in predictive ability in predicting out of sample relative to in sample (reported in the pooled column, as well as earlier tables). However, for all partitions and for all three models, the differences observed in predictive ability in the pooled column are preserved in the out-of-sample tests as well. For example, the percentage of bankruptcy firm-years in the top three deciles of the predicted hazard for the three models for nonrestatement years in the combined out of sample tests (column 4) is 81.75, 82.98, and 90.44 %, respectively. These percentages are considerably larger than the corresponding amounts for restatement years at 50.45, 64.86, and 69.37 %, respectively.

Discretion can affect predictive ability of the models in different ways. On the one hand, given a set of coefficients, the inputs to the prediction model may contain an error, which will affect the predicted values of the hazard. On the other hand, the optimal coefficients may be partition-specific. In other words, the coefficients appropriate to restatement firm-years differ from those of nonrestatement firm years. This can affect predictive ability in two ways. First, since the number of nonrestatement observations is much greater, their coefficients would dominate in a pooled regression, resulting in lower observed predictive ability for the unrestated group. Second, the inclusion of restated observations in the model will cause the coefficients estimated for the pooled sample to diverge from the optimal coefficients for unrestated observations, causing deterioration in predictive power of the model for this subgroup. This second effect could potentially mean that our analysis underestimates the effects of discretion on predictive power.

To test these effects, we estimate the coefficients separately for each partition. Note that, in doing so, we will be conducting an estimation that is not feasible in "real-time" because the firm-years that are subject to restatement are only known several years afterwards.

For all partitions and for all three models, the results, reported in fifth column of Table 10, are essentially the same as before. For example, when we estimate different coefficients for restated and nonrestated firm years separately, we find that 82.19, 82.98, and 90.7 % of bankrupt nonrestated years fall on the bottom three deciles of the accounting, market, and combined models. These percentages are considerably larger than the corresponding percentages for restatement years: 57.66, 64.86, and 70.27 %. Hence, differing coefficients across subsamples do not explain the deterioration in predictive ability discussed in the main body of the paper. The predictive ability of the accounting model slightly increases for the unrestated group, suggesting that the

Table 10 Out of sample tests and estimation by subgroup

	Accounting model					
	Pooled	Out of sample (1)	Out of sample (2)	Out of sample (3)	Subgroup	Stock exchange
Restated years	50.45	46.55	50.94	50.45	57.66	49.54
Non restated years	82.02	79.96	84.41	81.75	82.19	82.61
Low DACC	75.93	74.31	75.00	75.31	74.07	76.09
Medium DACC	81.82	80.10	83.19	80.61	82.89	62.68
High DACC	68.33	65.45	72.31	67.50	70.00	65.81
Average R&D = 0	80.00	79.26	79.75	79.01	79.75	79.67
Medium R&D	89.81	89.31	89.70	89.51	90.12	90.25
High R&D	63.03	59.68	68.42	61.34	73.95	61.98
Negative BTM	30.14	29.41	32.05	31.85	43.15	36.93
Low positive BTM	67.44	60.71	79.45	70.54	76.74	70.63
Medium positive BTM	79.07	79.72	79.14	78.71	84.97	79.89
High positive BTM	58.59	55.04	63.27	59.03	65.20	59.91
Loss years	70.75	69.72	71.27	70.87	70.81	68.53
Non loss years	77.62	77.61	76.40	77.35	77.62	81.89
	Market model					
	Pooled	Out of sample (1)	Out of sample (2)	Out of sample (3)	Subgroup	Stock exchange
Restated years	63.96	67.24	58.49	64.86	64.86	63.30
Non restated years	83.42	83.66	82.31	82.98	82.98	84.12
Low DACC	65.43	63.89	65.00	64.81	67.59	69.57
Medium DACC	86.23	86.65	86.04	85.83	86.23	87.55
High DACC	65.00	63.64	66.15	65.00	63.33	65.81
Average R&D = 0	82.73	84.94	80.00	82.48	82.73	82.06
Medium R&D	86.73	83.65	89.09	86.42	86.11	88.68
High R&D	73.11	74.19	71.93	73.11	72.27	68.60
Negative BTM	54.79	58.09	51.92	53.42	55.48	57.84
Low positive BTM	72.87	69.64	75.34	72.09	72.09	73.02
Medium positive BTM	81.75	81.14	83.09	81.93	81.22	82.25
High positive BTM	60.79	65.12	56.12	61.23	62.11	65.64
Loss years	61.30	62.44	60.48	60.85	60.85	65.45
Non loss years	83.43	80.10	87.58	83.43	83.43	84.96

Table 10 continued

	Combined model					
	Pooled	Out of sample (1)	Out of sample (2)	Out of sample (3)	Subgroup	Stock exchange
Restated years	68.47	68.97	67.92	69.37	70.27	73.39
Non restated years	90.88	91.21	90.54	90.44	90.70	90.51
Low DACC	75.31	76.39	75.00	75.93	75.93	79.19
Medium DACC	91.98	91.69	91.74	91.44	92.11	92.02
High DACC	75.00	69.09	78.46	75.83	73.33	76.07
Average R&D = 0	89.07	89.38	88.50	88.94	89.19	89.84
Medium R&D	92.59	91.82	92.12	92.59	94.14	94.34
High R&D	71.43	62.90	77.19	71.43	81.51	73.55
Negative BTM	55.14	52.94	56.41	53.08	58.90	55.40
Low positive BTM	73.64	67.86	79.45	74.42	82.17	75.40
Medium positive BTM	88.01	88.26	86.33	87.66	90.52	87.68
High positive BTM	73.57	73.64	76.53	74.45	78.41	74.89
Loss years	74.79	74.41	73.43	73.68	74.79	75.37
Non loss years	89.78	87.56	89.44	89.50	89.78	89.69

This table shows the percentage of firms in the top three deciles of the predicted probability of bankruptcy, where the probability of bankruptcy is estimated using different approaches. In the pooled estimation, coefficients are estimated for the full sample, as in Tables 3, 4, 5, 6, 7, 8. In the out of sample tests, the pooled sample is divided randomly into two subgroups, A and B. In the column labeled “Out of sample (1), we estimate the probability of bankruptcy for subsample B using the coefficients estimated for subsample A. We then rank the probability of bankruptcy within subsample B by year and subgroup and obtain the percentage of firms in the top three deciles of this estimated probability by subgroup. In “Out of sample (2),” we perform a similar analysis, switching subsamples A and B. In “Out of sample (3),” we pool the predicted probabilities of bankruptcy for subsamples A and B and rank them by year and subgroup. In the subgroup analysis, we estimate different sets of coefficients for each subgroup and also rank the predicted probability of bankruptcy within year and subgroup. In the stock exchange analysis, we estimate different sets of coefficients for each stock exchange

inclusion of restated observations affects the overall estimation of the model but has a very small effect on its predictive ability for unrestated observations.[14]

The NASDAQ and NYSE/AMEX samples exhibit statistically significant differences in the frequency of bankruptcy, the accounting and market variables and the partition variables (Table 2). To ensure that the observed differences in predictive power cannot be explained by differences in the optimal coefficients across stock exchanges, we re-run our main analysis using stock exchange specific coefficients. The last column in Table 10 presents the results from this analysis. In untabulated analysis we repeat the out-of-sample tests with stock exchange specific coefficients. All results are essentially the same as before.

[14] In untabulated analysis we use the “untainted” coefficients to predict the hazard for the entire sample. This doesn’t change our results.

In the above specifications, the baseline hazard is assumed to be constant across time. Bankruptcy rates are likely correlated, however, with fluctuations in economic activity. As a result, cross-sectional correlation of errors may be a concern in the above regressions, resulting in upward-biased standard errors. To circumvent this problem, and following Hillegeist et al. (2004), we use the overall frequency of bankruptcy in a given year to proxy for the baseline hazard. (This rate is calculated as the ratio between the number of bankruptcies and the total number of firms in the sample over the previous 12 months and is expressed as a percentage.) In unreported results, the annual bankruptcy rate is significant in all specifications, suggesting that the baseline hazard rate provides information that is incremental to the accounting and market variables.

In addition, we combine the market-based models into a Black–Scholes–Merton model of bankruptcy. We use the SAS code provided in the "Appendix" of Hillegeist et al. (2004) to estimate the BSM probability of bankruptcy, defined as the probability that the market value of assets is less than the face value of liabilities. Also following this study, the BSM probability of bankruptcy is then transformed into a score using the inverse logistic function. In unreported results, this variable is significant in all specifications.[15] In the basic specification, which merely includes the BSM score and the annual bankruptcy rate, the BSM score has a coefficient close to that reported in Table 5 of Hillegeist et al. (2004). The BSM model performs slightly worse than the market-based model. The accounting variables are still significant when the market variables are replaced by the BSM score, suggesting that the accounting information has incremental explanatory power with respect to this variable.

Lastly, the partitioning variables may be correlated with the probability of bankruptcy and this correlation may drive the result of lower predictive power for extreme values of the partition. We re-ran the analysis including the partitioning variables as explanatory variables in the base models. Our results are robust to this alternative specification.

In summary, none of the alternative specifications alters our conclusions regarding discretionary accruals, restatements, research and development intensity, book-to-market, and losses.

7 Concluding remarks

Our goal is to explore the effect of cross-sectional and time-series differences in discretion, unrecognized intangible assets, book-to market ratios, and incidence of losses on the predictive ability of financial ratios for bankruptcy. We find that all of our proxies for the exercise of discretion in financial reporting are associated with a significant deterioration in the predictive power of the accounting-based model. In addition, the presence of discretion impairs the predictive ability of not only the

[15] In the basic specification that merely includes the BSM score and the annual bankruptcy rate, both variables have magnitudes that are comparable to Hillegeist et al. (2004). In particular, the coefficient on the BSM score is 0.31 (vs. 0.27) and on the annual rate 0.43 (vs. 0.54)

accounting-based model but also the market-based and combined models. In other words, the total mix of information reflected in market-based variables, of which accounting data are a subset, does not offset or compensate for the effects of discretion.

We also find that the presence of intangible assets, as measured by R&D intensity, has a systematic effect on predictive ability. In particular, the predictive power of the accounting-based model is lower for firms with a high degree of R&D intensity.

We also examine the predictive power of the bankruptcy models across various categories of the book-to market ratio. Predictive power varies across book-to-market classes but not in a monotonic fashion. Firm-years with low to medium positive book-to-market ratios are most informative, consistent with more informative financial statements results when the book value of equity is closer to the market value of equity. Firm-years with high book-to-market ratios are next most informative, while the financial statements of firms with negative ratios of book-to-market are least informative. The findings are consistent with the contention that, when financial statements fail to recognize changes in asset values, either in the form of intangible assets or abandonment options, the predictive ability of financial ratios is impaired. These findings have potential implications for the use of the book-to-market ratios in other contexts as well.

We find that the incidence of a loss significantly increases the conditional probability of bankruptcy. However, we also find that the predictive power of the bankruptcy model for loss firm-years tends to be lower than for nonloss firm-years because of deterioration in the incremental explanatory power of the remaining variables. Perhaps surprisingly, market-based ratios do not compensate for the lower predictive ability of financial ratios in loss years and instead reflect substantially less information useful for predicting bankruptcy.

Finally, we conduct time-series analysis to improve our understanding of factors influencing the decline in the accounting model's predictive power over time. We find that there is a significant time trend in the frequency of restatements, of larger magnitudes of discretionary accruals, of greater R&D intensity, of book-to-market ratios that are further from one, and of losses. These variables are individually significant in explaining differences in predictive ability over time. However, because of high correlation with each other, it is difficult to isolate individual, incremental effects.

In the cross-sectional context, in most cases, the market model also exhibited lower predictive power for the same categories of firm-years as the accounting model. However, unlike the accounting model, the market model exhibits no declining time trend, and differences in predictive power over time are uncorrelated with our partitioning variables. These findings suggest that the changes in financial reporting attributes we document contribute to less informative financial ratios, as assessed by bankruptcy prediction. They do not contribute to less informative market variables over time. Furthermore, the findings that the combined model exhibits a declining time trend in predictive power and that this is associated with our partitioning variables indicate that the market variables included in our market and combined models did not compensate for the loss of information over time.

Acknowledgments The authors thank David Aboody, Judson Caskey, Jack Hughes, Bruce Miller, Richard Saouma, Lakshmanan Shivakumar, Brett Trueman, Irem Tuna, Florin Vasvari, and seminar participants at UCLA and London Business School for helpful comments on an earlier version of this paper. We also thank Jim Ohlson and an anonymous referee for many helpful suggestions. We greatly appreciate the generosity of Tyler Shumway, Sudheer Chava, and Robert Jarrow in making data available to us for this project.

Appendix

Variable definitions and data sources (Tables 11, 12).

Table 11 Panel A: variable definitions

Variable	Description
BTM	Ratio between the book value of equity (Compustat data 216) and market capitalization at fiscal year end (Compustat data25*data199)
CUMPCT	Cumulative percentage of year of bankruptcy observations in the top three deciles of the hazard rate
DACC	Difference between total current accruals and nondiscretionary accruals (following Dechow et al. 1995, nondiscretionary accruals are obtained by running a cross-sectional regression of current accruals on change in sales, adjusted by the change in receivables, with both the independent and dependent variables scaled by lagged total assets)
DREST	Dummy variable equal to one if the firm year was restated. (This variable is obtained from the sources described in "Appendix".)
ETL	Net income before interest, taxes, depreciation, depletion, and amortization divided by total liabilities, both short- and long term (Compustat data13/data181)
LERET	Prior year's security returns, where security returns are calculated over the 12 months ending with the third month after the end of the fiscal year
LRSIZE	Logarithm of the market capitalization as of the end of the third month after the end of the fiscal year, divided by the market capitalization of the market index of NYSE, AMEX, and NASDAQ firms
LSIGMA	Standard deviation of the residual return from a regression of the security's monthly return on the return of the market portfolio (the return for the 12 months ending with the third month of the fiscal year is used in this regression)
LTA	Ratio between total liabilities and total assets (Compustat data 181/data 6)
MBTM	Percentage of firms in year t in deciles 5 through 8 of BTM
MDREST	Percentage of sample firms that restated their previously issued financial statements for year t
MHIGHDACC	Percentage of firms in year t whose discretionary accruals represent more than 10 % of lagged assets
MHIGHRD	Percentage of firms in year t whose R&D expenses represent more than 5 % of sales
MNROAI	Percentage of loss firms in year t
NROAI	Dummy variable equal to 1 if the return on assets (ROA) is negative
RDSALES	R&D expenses divided by sales (Compustat data46/data12)
ROA	Return on assets, defined as earnings before interest adjusted for interest and income tax and scaled by lagged assets (Compustat data172+data15*(1-taxrate))/lagged data6)
T	Time trend

Table 12 Panel B: data sources for the restatement variable

Data source	Period		Content	Description
	Begin	End		
GAO restatement database	1997	June 2002	Restatements	This database was collected by the GAO, based on public sources and is available on the GAO website (http://www.gao.gov/new.items/d03395r.pdf). It contains financial statement restatements announced from January 1, 1997, through June 30, 2002. We hand collected the periods affected by the restatement, by reading the respective press releases
Huron database	1997	2003	Restatements	This is a proprietary database generously supplied by the Huron Consulting Group. It includes restatements announced between 1997 and 2003 and the respective manipulation period
Hand collected data from SEC website	1995	2003	AAER	We hand collected the Accounting and Auditing Enforcement Releases (AAERs) for 1995 through 2003 from the SEC website and read each AAER to determine the manipulation period. (http://www.sec.gov/divisions/enforce/enforceactions.shtml)
Bonner, Palmrose and Young (1998)	1982	1995	AAER	This database, generously supplied by Bonner, Palmrose, and Young, includes AAERs for the 1982 through 1995, as well as the manipulation period
Woodruff-Sawyer	1980	2000	Class action securities lawsuits	This database, provided by Woodruff-Sawyer & Co., includes firms sued by shareholders for accounting improprieties during the period 1980 through 2000. It contains virtually the entire population of federal shareholder lawsuits filed in this period. The sources for this database are the Securities Class Action Alert newsletters, the Securities Class Action Clearinghouse of Stanford Law School, press releases and wire service reports, the IPO Reporter newsletter, Moody's Corporation Data System, and various law firms and claims administration services

References

Altman, E. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance, 23*, 589–609.

Barth, M., Clement, M., Foster, G., & Kasznik, R. (1998). Brand values and capital market valuation. *Review of Accounting Studies, 3*, 41–68.

Basu, S. (1997). The conservatism principle and the asymmetric timeliness of earnings. *Journal of Accounting and Economics, 24*, 3–37.

Beaver, W. H. (1965). Financial ratios as predictors of failure. Unpublished Ph.D thesis. Graduate School of Business, University of Chicago.

Beaver, W. H. (1966). Financial ratios as predictors of failure. *Journal of Accounting Research, 4*, 71–102.

Beaver, W. H. (2002). Perspectives on recent capital market research. *The Accounting Review, 77*, 453–474.

Beaver, W. H., & Engel, E. (1996). Discretionary behavior with respect to allowances for loan losses and the behavior of security prices. *Journal of Accounting and Economics, 22*, 177–206.

Beaver, W. H., McNichols, M., & Rhie, J. (2005). Have financial statements become less informative? Evidence from the ability of financial ratios to predict bankruptcy. *Review of Accounting Studies, 10*, 93–122.

Bharath, S., & Shumway, T. (2008). Forecasting default with the Merton distance to default model. *Review of Financial Studies, 21*, 1339–1369.

Bonner, S., Palmrose, Z., & Young, S. (1998). Fraud type and auditor litigation: An analysis of SEC accounting and auditing enforcement releases. *The Accounting Review, 73*, 503–552.

Bradshaw, M., & Sloan, R. (2002). GAAP versus the street: An empirical assessment of two alternative definitions of earnings. *Journal of Accounting Research, 40*, 41–66.

Brown, S., Lo, K., & Lys, T. (1999). Use of R2 in accounting research: Measuring changes in value relevance over the last four decades. *Journal of Accounting and Economics, 28*, 83–115.

Campbell, J., Hilscher, J., & Szilagyi, J. (2008). In search of distress risk. *Journal of Finance, 63*, 2899–2939.

Chava, S., & Jarrow, R. (2004). Bankruptcy prediction with industry effects. *Review of Finance, 8*, 537–569.

Collins, D., Maydew, E., & Weiss, I. (1997). Changes in the value-relevance of earnings and book values over the past forty years. *Journal of Accounting and Economics, 24*, 39–67.

Collins, D., Pincus, M., & Xie, H. (1999). Equity valuation and negative earnings: The role of book value of equity. *The Accounting Review, 74*, 29–61.

Dechow, P., & Schrand, C. (2004). *Earnings quality*. Charlottesville, VA: The Research Foundation of the CFA Institute.

Dechow, P., Sloan, R., & Sweeney, A. (1995). Detecting earnings management. *The Accounting Review, 70*, 193–225.

Easton, P., Monahan, S., & Vasvari, F. (2009). Initial evidence on the role of accounting earnings in the bond market. *Journal of Accounting Research, 47*, 721–766.

Financial Accounting Standards Board (FASB). (2001). Accounting for the impairment or disposal of long-lived assets. *FASB statement no. 144.* Stamford, CT: FASB.

Francis, J., & Schipper, K. (1999). Have financial statements lost their relevance? *Journal of Accounting Research, 37*, 319–352.

Francis, J., Schipper, K., & Vincent, L. (2002). Expanded disclosures and the increased usefulness of earnings announcements. *The Accounting Review, 77*, 515–546.

Francis, J., LaFond, R., Olsson, P., & Schipper, K. (2004). Cost of equity and earnings attributes. *The Accounting Review, 79*, 967–1010.

Francis, J., LaFond, R., Olsson, P., & Schipper, K. (2005). The market pricing of accruals quality. *Journal of Accounting and Economics, 39*, 295–327.

Franzen, L., Rodgers, K., & Simin, T. (2007). Measuring distress risk: The effect of R&D intensity. *Journal of Finance, 57*, 2931–2967.

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

Graham, B., & Dodd, D. (1934). *Security analysis*. New York, NY: McGraw Hill.

Hayn, C. (1995). The information content of losses. *Journal of Accounting and Economics, 20*, 125–153.
Hillegeist, S., Keating, E., Cram, D., & Lundstedt, K. (2004). Assessing the probability of bankruptcy. *Review of Accounting Studies, 9*, 1573–7136.
Hribar, P., & Nichols, D. C. (2007). The use of unsigned earnings quality measures in tests of earnings management. *Journal of Accounting Research, 45*(5), 1017–1053.
Kasznik, R. (1999). On the association between voluntary disclosure and earnings management. *Journal of Accounting Research, 37*, 57–81.
Landsman, W., & Maydew, E. (2002). Has the information content of quarterly earnings announcements declined in the past three decades? *Journal of Accounting Research, 40*, 797–808.
Lev, B. (2001). *Intangibles: Management, measurement, and reporting*. Washington, D.C.: Brookings Institution Press.
Lev, B., & Sougiannis, T. (1996). The capitalization, amortization, and value-relevance of R&D. *Journal of Accounting and Economics, 21*, 107–138.
Lev, B., & Zarowin, P. (1999). The boundaries of financial reporting and how to extend them. *Journal of Accounting Research, 37*, 353–386.
McNichols, M. (2000). Research design issues in earnings management studies. *Journal of Accounting and Public Policy, 19*(2000), 313–345.
Merton, R. (1974). On the pricing of corporate debt: The risk structure of interest rates. *Journal of Finance, 29*, 449–470.
Ohlson, J. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research, 18*, 109–131.
Shumway, T. (2001). Forecasting bankruptcy more accurately: A simple hazard model. *Journal of Business, 74*, 101–124.
Wahlen, J. (1994). The nature of information in commercial bank loan loss disclosures. *The Accounting Review, 69*, 455–478.
Watts, R., & Zimmerman, J. (1990). Positive accounting theory: A ten year perspective. *The Accounting Review, 65*, 131–156.