On the Role of Industry-Level Structural Constraints and the Timing of Accounting Reports in Bankruptcy Predictions

Wan-Chien Chiu[a], Juan Ignacio Peña[a], and Chih-Wei Wang [a]

This version: July 2011

Abstract

This study uses a hazard model with data on 3392 corporate bankruptcies by U.S. public companies during 1983–2008 to determine the effect of industry-based structural constraints on bankruptcy predictions. The probability of bankruptcy is significantly higher for firms in highly concentrated industries and with relatively stronger customer dependency. Most bankruptcy predictions reflect the variation of a firm's characteristics relative to its industry, but industry-specific characteristics have negligible impacts. The investigation also includes a comparison of the relative performance of accounting and market-based variables, in terms of both in-sample fit and out-of-sample forecasting accuracy. For yearly data, the best model includes both accounting and market-based variables. However, for monthly market data and quarterly accounting reports, the best model features only market data. The usefulness of accounting measures in bankruptcy prediction models thus may be contingent on sampling frequency.

Keywords: Bankruptcy Prediction, Industry Effects, Market and Accounting Variables.

JEL classification codes: G32, G33, L00.

1. Introduction

Two factors likely determine the forecasting accuracy of corporate bankruptcy hazard rate models. Specifically, we are interested in the effects of the intensity of competition and the degree of connection with customers and suppliers in a given industry, as well as potential distortions in accounting information across different time spans (e.g., quarterly versus yearly reports). If accounting signals in financial reports are less informative at the quarterly than at the yearly level, the latter should provide more accurate bankruptcy forecasting. Moreover, industry structural constraints should affect firms' operating strategies, financial structure decisions, and profitability, and therefore their bankruptcy likelihood. Maksimovic and Phillips (1998) indicate that bankruptcy reflects industry demand conditions, and Chava and Jarrow (2004) provide empirical support but use only dummy variables, without discussing specific factors. We instead determine just how bankruptcy might be influenced by a set of structural constraints within each industry.

Growing research based on hazard rate models compares the relative merits of accounting and market-driven variables for bankruptcy forecasting. Shumway (2001) finds that combining the two can produce accurate forecasts; Chava and Jarrow (2004) instead suggest that monthly and quarterly observations significantly improve the forecasting performance of both accounting and market-based models, compared with yearly observations. This result is consistent with the idea that timely information provision is particularly useful in forecasting bankruptcy. However, the contribution of accounting variables grows negligible in the presence of market variables; we therefore consider whether previous results still hold when we rely on an extended database and a more recent period.[1]

---

[1] Our database contains 3392 bankruptcies between 1983 and 2008; Chava and Jarrow's (2004) database featured 1461 bankruptcies between 1962 and 1999.

To assess the effect of an industry's structural constraints on bankruptcy predictions, we consider corporate bankruptcies by U.S. public companies during the period from 1983 to 2008. Our in-sample tests include the full sample; for our out-of-sample tests, we use data from 1983–1999 to fit the hazard model, then adopt the estimated coefficients to forecast bankruptcies during 2000–2008. Both the in- and out-of-sample tests indicate that the probability of bankruptcy is significantly higher for firms in highly concentrated industries, compared with similar firms in industries with lower degrees of concentration. We also find that firms that depend more on customers have higher bankruptcy probabilities than firms with lower dependency levels. Both findings highlight the risks inherent in low flexibility, whether in the kind of product offered to the market (i.e., concentrated industries tend to produce specialized products) or the customers served. We also find that our ability to predict bankruptcy depends almost completely on the variation of a firm's characteristics from its industry; industry-specific characteristics instead have a negligible impact on this accuracy level.

Regarding the relative performance of accounting and market variables, both in their in-sample fit and out-of-sample forecasting accuracy, we find that models based on yearly data have 25% more explanatory power on average than models based on monthly and quarterly data. When using yearly data, the best model (in and out-of-sample) includes both accounting and market-based variables, but if we employ monthly market data and quarterly accounting data, the best model only includes market data. On the one hand, this finding suggests potential distortions in quarterly accounting signals, for which financial reporting flexibility is much wider than it is at the yearly level. This flexibility may allow managers to produce quarterly reports that fail to reflect the underlying economic fundamentals, such that the usefulness of accounting measures in bankruptcy prediction models may be

contingent on sampling frequency. On the other hand, an alternative explanation is based on the hypotheses that market prices reflect all publicly available information regarding bankruptcy because for monthly data, market variables are updated more quickly than are the quarterly accounting variables. For yearly data, updating of both market and accounting variables occurs at the same frequency.

In turn, we contribute to empirical literature on bankruptcy prediction in several ways. We confirm the importance of accounting for the degree of concentration in the industry, the dependency of customers, and the firm's characteristics relative to its industry when forecasting bankruptcies. We note the relative explanatory power of market and accounting variables in hazard models and document differences when using yearly or quarterly accounting data; we recommend the former in terms of its forecasting ability.

The rest of this article is organized as follows: In Section 2, we relate relevant literature to our hypotheses about the effect of industry constraints in bankruptcy prediction. Then in Section 3, we define our data. Section 4 contains the bankruptcy prediction model and a discussion of the empirical results, both in-sample and out-of-sample, and Section 5 offers the model of within- and across-industry variables. We conclude in Section 6.

2. Bankruptcy Prediction with Internal and External Industry Constraints

Previous literature has noted the correlation between bankruptcy likelihood and a firm's capital structure; see Titman and Wessels (1988). Although Chava and Jarrow (2004) provide empirical evidence of the importance of industry effects, as far as we know, the actual influence of industry characteristics on bankruptcy likelihood has not been addressed. To investigate this point, we measure structural constraints both within and across industries; accordingly, we propose two hypotheses with respect to internal and external industry constraints.

2.1. Internal Industry Constraints

Industrial organization literature reveals that the structure of product markets affects managers' operating decisions (Brander and Lewis, 1986). The consequences of operating decisions may affect the risk of a firm's cash flows and therefore stock returns. The structure of product markets also determines firms' capital structure choices and therefore their leverage risk. Firms with a high leverage risk are more likely to file for bankruptcy, given their high level of debt compared with their equity.

Opler and Titman (1994) argue that customers are more reluctant to purchase products from a distressed firm with specialized products that may need servicing. Because firms in highly concentrated industries are more likely to produce specialized products, they become more vulnerable to financial distress. In concentrated industries and weak economies, highly leveraged firms thus tend to lose more market share than do conservatively financed competitors, which implies even greater losses for financially distressed firms in concentrated industries.

In line with risk-based return considerations, Hou and Robinson (2006) posit that lower expected returns for firms in highly concentrated industries result because they face fewer risks. A conventional view suggests that firms that bear more risk need to provide higher returns as compensation, though Campbell, Hilscher, and Szilagyi (2008) empirically contradict that claim. Financially distressed stocks deliver anomalously lower returns but much higher standard deviations, market betas, and loadings on value and small-cap risk factors than do stocks with low default risk. In other words, firms with lower returns may have higher, not lower, risks. In turn, the relationship between the degree of industry concentration and firm risk remains an open question for empirical investigation.

Overall, firms in highly concentrated industries may be more likely to file for bankruptcy for two reasons. First, they tend to become more leveraged than firms in

less concentrated industries. Second, firms in concentrated industries provide low average stock returns but suffer from relatively high volatility, as suggested by Campbell et al. (2008). We hypothesize in turn that

Hypothesis 1: The higher an industry's concentration ratio, the higher the incidence of bankruptcy among its firms.

To proxy for the degree of concentration in an industry, we employ its concentration ratio, or the inverse of the intensity of industrial competition. We use a Herfindahl index to measure the degree of industry concentration in industry j as follows:

$$\text{Herfindahl}_j = \sum_{i=1}^{I} s_{ij}^2 \quad (1)$$

where s_{ij} is the market share of firm i in industry j. Small values imply that the market contains many competing firms; large values imply that few large firms dominate the market. To construct the Herfindahl measure, we use the entire distribution of industry market share (i.e., net sales) information to assess the degree of industry concentration. For this study, we compute the Herfindahl index for each year (yearly data set) or month (monthly data set) and for each industry. Because the industry's internal constraint is the competition between producers in the same market seeking the same business, the inverse Herfindahl index offers a good proxy of industry internal constraints.

2.2 External Industry Constraints

Hertzel, Li, Officer, and Rodgers (2008) argue that financial distress at one firm links not only to the product market (internal industry) but also to external industries, because firms connect with customers and suppliers in other (external) industries. We thus consider how external industry constraints may influence bankruptcies, especially in relation to customers and suppliers. Although a firm's customers and suppliers might belong to the same industry as the focal firm, our hypothesis refers

only to those in other industries.

The trade-off theory of capital structure implies that the actions of distressed firms' customers and suppliers increase indirect costs and can worsen financial distress or even result in bankruptcies. On the one hand, when customers confront a supplier in distress, they likely impose costs and find new suppliers to replace it, due to their concerns about product quality, reduced warranty value, continuity of supply, and serviceability. On the other hand, when suppliers find a customer in financial distress, they may impose costs by failing to supply trade credit, backing away from long-term contracts, or delaying shipments. This viewpoint emphasizes the potential importance of indirect costs, such that we anticipate that firms that depend more on their customers and suppliers suffer higher indirect costs when in financial distress. In more general terms, a firm with greater dependence on its customers and suppliers has more chances to interact with them in all states of nature, including in distress states. However, the magnitude of the indirect costs is difficult to estimate.

Burt (1988) offers that the competitive advantages of firms increase with the weakness of the connections among customers and suppliers in their industry. Burt (1983, 1988), Talmud (1994), and Yasuda (2005) further suggest that both customer and supplier constraints negatively affect a firm's economic performance. Moreover, in an industry with low bargaining power (i.e., higher customer and supplier dependency or constraints), corporate strategies depend more on customers and suppliers, which increases the costs of gathering resources and distributing products.

Theoretical and empirical work on capital structure also has addressed economic links to customers and suppliers during financial distress episodes. Regarding non-financial stakeholders' (NFS) bargaining power, Kale and Shahrur (2007) find that a firm tends to have higher debt ratios when its customers and suppliers take strong bargaining positions. Firms with more debt have more chances to file for

7

bankruptcy. Yet Sarig (1998) shows that when a firm worries about suppliers threatening to curtail its specialized factors of production, it will limit its leverage both before and after contracting, to prevent potential NFS hold-up behavior. Lowering the debt ratio should reduce the likelihood that unstable supply will lead to the firm's liquidation. However, we find no clear conclusion about the effect of suppliers' bargaining power on firms' leverage, nor do we have any compelling arguments about whether suppliers' constraints raise the firm's bankruptcy chances or not.

Customers' constraints often arise when one specific supplier is closely connected to its buyers. The supplier suffers a weaker bargaining power position and also suffers more losses when its buyers cannot pay, especially during financial distress periods. In a booming economy, a supplier that faces high customer constraints earns lower profits; in a bust economy, it cannot collect payments from its customers. Thus both forces imply that firms with more customer constraints are more likely to file for bankruptcy.

On the other side, high supplier constraints indicate that the firm has a stronger connection to its suppliers. It again takes a weaker bargaining power position but does not suffer the same potential monetary losses, because it is the customer, obligated to pay its suppliers, not receive payments. Therefore, a firm facing strong suppliers' constraints may not go bankrupt as easily as a firm with high customers' constraints. Thus we formulate our second hypothesis:

Hypothesis 2: The stronger a firm's customers' constraints, the higher its incidence of bankruptcy. The relationship between a firm's suppliers' constraints and its bankruptcy is unclear.

We use a measure of external industry constraints proposed by Burt, Guilarte, Raider, and Yasuda (2002), which begins with a consideration of resource dependence, or the extent to which producers in a market depend on another market to buy or sell,

directly or indirectly. They simultaneously consider constraints derived from customer and supplier markets. To investigate the impact of customers' and suppliers' constraints on bankruptcy prediction specifically, and following Shih's (2007) suggestion, we partition Burt et al.'s (2002) combined measure of external industry constraints into customers' constraints and suppliers' constraints. Burt et al. (2002) measure external constraints on industry i as the sum of transaction-specific constraints on the market, using the value of commodities sold to industry j from industry i. We denote $C_{i,B}$ and $C_{i,S}$ as customers' and suppliers' constraints, respectively, in industry i, then calculate them using the following method:

$$C_{i,type} = \sum_j c_{ij,type} \ , \quad i \neq j \ , \ type = B \ or \ S, \quad (2)$$

where $\quad c_{ij,type} = \left(p_{ij,type} + \sum_q p_{iq,type} \cdot p_{qj,type} \right)^2 \cdot H_j \ , \quad i \neq q \neq j \ ; \quad p_{iq,B} = \dfrac{z_{iq}}{\sum\limits_j z_{ij}} \ ; \quad$ and

$p_{iq,S} = \dfrac{z_{qi}}{\sum\limits_j z_{ji}} , \ i \neq j$. Moreover, H_j is the concentration ratio of industry j, computed by the Herfindahl index; and z_{ij} is the dollar worth of the commodities sold to industry j from industry i. To provide the intuition underlying this measure, we use customers' constraints as an example: The squared term in Equation (2) is the degree of direct and indirect dependence of industry i on market j, measured by p_{ij}, or the proportion of industry i's sales that occur directly to market j, plus the proportion of industry i's sales that indirectly involve market j through market q. Then H_j is the extent to which the buyer market j is oligopolistic. Finally, $C_{i,B}$ and $C_{i,S}$ vary between 0 and 1, and the later scenario implies the maximum constraint. A firm in an industry with $C_{i,B}$ ($C_{i,S}$) equal to 0 has maximal bargaining power as a supplier (customer).

3. Data

For our empirical study, we select all firms traded on NYSE/AMEX/NASDAQ

and those with available data in the Center for Research in Security Prices (CRSP) and COMPUSTAT databases to construct explanatory variables. To discriminate bankrupt firms from non-bankrupt ones, we adopt a broad definition of bankruptcy that includes firms delisted due to bankruptcy, liquidation, or poor performance (Dichev, 1998; Brockman and Turtle 2003). Bankrupt firms in our sample meet two conditions: They were listed in the CRSP database but then were delisted during our sample period, and the delisting code reported by CRSP equals 400 or 550–585.[2] We identify 3392 bankruptcies according to these criteria, more than any other published study to the best of our knowledge.

3.1. Descriptive Statistics

Figure 1 depicts bankruptcies as a percentage of active companies per year in our sample.[3] The general shape of this figure reflects national recessions, as defined by the National Bureau of Economic Research (NBER),[4] which has reported three recessions since 1983. The early 1990s recession related to the 1990 oil price shock, the debt accumulation of the 1980s, new banking regulations following the savings and loan crisis, and growing consumer pessimism. In Figure 1, we find a notable peak from 1990 to 1992. The second recession, in the early 2000s, combined the collapse of the speculative dot-com bubble and the September 11 attacks. Again, in Figure 1, we find a relative higher bankruptcy incidence in 2000–2002. The most recent, ―Great Recession,‖ runs from December 2007 to June 2009. Although we cannot observe the full effects of this last episode after 2008 due to data limitations, we find a big increase in bankruptcies in 2008, consistent with economic facts.

---

[2] The Appendix provides definitions of the CRSP delisting codes.
[3] Our sample period starts in 1983, but Figure 1 does not contain the percentage of bankruptcy in 1983, because our bankruptcy prediction model relies on one-year ahead data as independent variables. Therefore, our actual sample only includes bankruptcy data from 1984.
[4] The NBER defines a recession as ―a significant decline in economic activity spread across the economy, lasting more than a few months, normally visible in real gross domestic product (GDP), real income, employment, industrial production, and wholesale-retail sales.‖

[INSERT FIGURE 1 HERE]

Figure 2 instead segments bankruptcies by stock exchange listing. More than 80% of the firms in our sample that filed for bankruptcy were listed on NASDAQ. The remaining percentage is almost evenly spread between the NYSE and AMEX.[5]

[INSERT FIGURE 2 HERE]

3.2. Description of Input–Output Table

To measure external industry constraints—namely, customers' and suppliers' constraints—we use summarized versions of the input–output tables published by the Bureau of Economic Analysis (BEA). Specifically, the ―Use Table of Input-Output Accounts for the U.S. Economy" reports a matrix with the values of commodity flows between each pair of industries. However, the industry classification in this table changed after 1998, from Standard Industrial Classifications (SIC) to the North American Industry Classification System (NAICS). To be consistent, we divide our sample into two periods: 1983–1997 (SIC) and 1998–2008 (NAICS). Also, the table was published every five years before 1997, after which annual updates were published. For consistency, we thus use data reported in the 1987 ―Benchmark Use Table" to represent the period 1983–1989, and then use the data in the 1992 ―Benchmark Use Table" for the period 1990–1997. After 1998, we use annual data for each year through 1998–2008. Although the input–output tables classify the industries based on SIC or NAICS codes, we only observe the input–output (IO) industry code to classify industries into tables. To group firms in the CRSP database into different industries based on IO code, we use an SIC–IO codes conversion table,[6] as well as a document relating NAICS codes to IO codes.[7] Finally, all firms before 1997 are

---

[5] Shumway (2001) considers the NYSE and AMEX, where bankruptcies happen with a much lower probability than in the NASDAQ.

[6] The SIC–IO code conversion table is available on Joseph P.H. Fan's personal website.

[7] This document, available on the BEA website, is ―A Document for 2002 Benchmark Input-Output Accounts: Summary Make Table and Use Table before Redefinitions."

grouped into eight industries; after 1998 they span fourteen industries in our sample set.[8]

Table 1 contains an overview of bankruptcies by IO codes. According to Panel A, before 1997, most bankruptcies occurred in manufacturing (37%) or services (19%). As Panel B of Table 1 indicates, after 1998, almost 39% of the total bankruptcies happened in the manufacturing sector, followed by the information industry (15%) and then finance, insurance, real estate, rental, and leasing, which account for 12% of the total bankruptcies registered during this period.

[INSERT TABLE 1 HERE]

3.3. Definition of Independent Variables

We use a dynamic logit regression method to analyze bankruptcy prediction (see Shumway, 2001) with five explanatory variables:

(1) Excess return (EXRET), or the monthly log excess return on each firm' stock return relative to the value-weighted CRSP NYSE/AMEX/NASDAQ index return. Low past excess returns should increase the chance of bankruptcy, because the decrease in equity value increases leverage and therefore default probabilities.

(2) Relative size (RSIZ), which is the logarithm of each firm's market equity value (outstanding shares × stock prices) divided by the total NYSE/AMEX/NASDAQ market equity value. The lower the relative size, the higher the default probability, such that we expect a negative relationship between RSIZ and bankruptcy likelihood.

---

[8] The initial eight industries were (1) agriculture, forestry, fishing and hunting; (2) mining; (3) construction; (4) manufacturing; (5) transportation, communication and utilities; (6) trade; (7) finance, insurance, real estate, rental, and leasing; and (8) services. The fourteen industries instead are specified as (1) agriculture, forestry; fishing and hunting; (2) mining; (3) utilities; (4) construction; (5) manufacturing; (6) wholesale trade; (7) retail trade; (8) transportation and warehousing; (9) information; (10) finance, insurance, real estate, rental, and leasing; (11) professional and business services; (12) educational services, health care, and social assistance; (13) arts, entertainment, recreation, accommodation, and food services; and (14) other services, except government.

(3) Volatility of each firm's stock returns (SIGMA)—a sample standard deviation using the last 60 observable daily returns.[9] It measures the probability of the firm's asset values being below the default threshold, so we expect that higher SIGMA indicates a higher probability of bankruptcy.

(4) NI/TA, or the firm's net income divided by the adjusted total assets (see (5) below). This ratio typically represents a firm's profitability, and we expect a negative effect on bankruptcy.

(5) TL/TA, or the firm's total liabilities divided by the adjusted total assets. As a proxy of leverage, a firm with high TL/TA tends to have high probability of bankruptcy. We use adjusted total assets, instead of the raw total assets data from COMPUSTAT, to prevent potential errors of measurement in the assets' book values (TA) or generate outliers when used as denominators in calculating financial ratios. Thus we adjust the total assets following the procedure suggested by Cohen, Polk, and Vuolteenaho (2003). The adjusted total assets are the sum of 10% of the difference between market and book equity and the book value of total assets (see Campbell et al., 2008):

$$Total\ Assets\left(adjusted\right)_{i,t} = TA_{i,t} + 0.1\left(ME_{i,t} - BE_{i,t}\right)\ (3)$$

In addition to standard profitability and leverage variables, liquidity may determine bankruptcy. Firms often default because of their inability to pay their financial obligations on time due to the imbalance between their cash inflows and outflows (see Laitinen and Laitinen, 1998). To the extent that managers opportunistically manipulate accruals, earnings (which depend significantly on net

---

[9] The stock's volatility measure comes from Chava and Jarrow (2004) and differs from the method offered by Shumway (2001), who calculates each firm's sigma for year t by regressing each stock's monthly returns in year $t-1$ on the value-weighted NYSE/AMEX/NASDAQ index return for the same year. Sigma is the standard deviation of the residual of this regression.

income) offer a less reliable measure, and cash flow–based variables might be preferable.[10] Thus we add a measure of liquidity in the form of the ratio of a company's cash and short-term assets to its adjusted total assets (CASH/TA). An increasing CASH/TA should indicate a smaller chance of bankruptcy.

We also include the market-to-book ratio (MB), or the market value of equity (ME) divided by book value of equity (BE).[11] The book value of equity is as defined by Davis, Fama, and French (2000); following Cohen, Polk, and Vuolteenaho (2003), we replace negative BE values with small positive values of $1 to ensure that the market-to-book ratios are in the right, not the left, tail of the distribution. The MB variable captures the relative value of the firm's equity from stockholders' and accountants' viewpoint. Firms close to bankruptcy often have suffered eroded book value of their equity, driving up their market-to-book ratio; however, the stock market often discounts the market value of equity, given traders' negative predictions about those firms, driving down the market-to-book ratio. Thus, though MB has the potential to predict bankruptcy, it is unclear how it relates empirically to bankruptcy. For the time being, we anticipate that increasing MB results in increasing bankruptcy likelihood, in line with empirical evidence from Campbell et al. (2008).

Finally, to test for the effect of industrial structural constraints, we consider the internal industry constraint, measured by our Herfindahl index (H), and external industry constraints, including buyers' ($C_{i,B}$) and suppliers' ($C_{i,S}$) constraints, as we

---

[10] Cash flow might be a bankruptcy predictor: ―Many financial analysts regard operating cash flow as a better gauge of corporate financial performance than net income, since it is less subject to distortion from differing accounting practices" *(Chemical Week,* May 8, 1991, p. 28).

[11] BE is the stockholders' book equity, plus balance sheet deferred taxes and investment tax credit (if available), minus the book value of preferred stock. We measure stockholders' equity as the book value of common equity plus the par value of preferred stock, or the book value of assets minus total liabilities (in that order).

reviewed previously.

3.4. *Yearly vs. Quarterly and Monthly Observation Intervals*

Chava and Jarrow (2004) posit that monthly observation intervals can improve bankruptcy prediction, compared with yearly frequencies, because more frequent accounting reports (quarterly) provide more updated information about the company's financial health. However, considering the potential for earnings management in accounting reports, we wonder if quarterly accounting reports may be more accurate. That is, companies might manage their earnings to avoid reporting losses or earnings decreases or to meet analysts' expectations at the end of each year.[12] Although earnings management could be used both quarterly and annually, prior research suggests that the extensiveness of such ―massaging" of accounting information is far greater in quarterly reports, for three main reasons.

First, a firm can increase its quarterly earnings indirectly through accounting methods, such as by increasing accruals, or directly by postponing expenses to future quarters (Bernhardt and Campello, 2007). Matsunaga and Park (2001) show that CEO bonuses provide managers an incentive to meet the earnings reported for the same quarter of the prior year, for example. Second, errors in a previous quarterly report must be corrected in the annual report,[13] and on average, these error corrections are negative, such that original quarterly reported earnings appear overstated (Kinney and McDaniel, 1989). Third, accounting regulations do not mandate independent audits of quarterly reports, so managers may find it easier to fudge interim quarterly earnings, compared with earnings in the fourth quarter. Auditing provides a mechanism for ensuring that companies provide accurate information, and auditors should warn

---

[12] See, for example, Hayn (1995); Burgstahler and Dichev (1997); Degeorge, Patel, and Zeckhauser (1999); Dechow, Richardson, and Tuna (2003); and Roychowdhury (2006).
[13] The SEC's Accounting Series Release (ASR) No. 177 mandates that corrections of errors in quarterly earnings reports of public firms be uncovered in the annual report.

investors if a company appears likely to file for bankruptcy. Thus audited annual earnings likely are more credible and accurate than interim quarterly earnings with regard to the company's true financial health.

Existing literature investigates whether capital market participants may be fooled by earnings management; one article suggests that investors cannot see through cases of earnings management, and investors may fail to price firms properly to include earnings shocks caused by accruals (Healy and Wahlen, 1999).[14] In addition, evidence indicates that firms that manage to meet or beat their earnings expectations enjoy an average quarterly return almost 3% higher than that of peers that fail to do so, even if investors discount the return after they notice earnings management practices (Bartov, Givoly, and Hayn, 2002). One study pays particular attention to the comparison of stock performance by firms that issue accurate annual reports and those that fudge quarterly reports. Myers, Myers, and Skinner (2007) document that firms that report earnings strings of at least 20 quarters (five years) with consecutive non-decreases enjoy abnormal returns of more than 20% per year during the first five years, much larger than the returns of firms that report five years of consecutive increases in annual (but not quarterly) reports.

Thus market participants appear to incorrectly value firm performance due to earnings management efforts, especially when companies manipulate their quarterly reports. The market also incorrectly interprets this accounting information. For a bankruptcy model then, market variables may not reflect firms' true financial situation. To conclude that using quarterly accounting data is better than using yearly data thus is debatable and open to challenge from empirical evidence.

Moreover, we cannot confirm if the superiority of monthly data over yearly data,

---

[14] Kinney and McDaniel (1989) show that the market may be misled by misstated quarterly earnings reports, based on the significant difference in the price performance of firms with over- versus understatements.

as suggested by Chava and Jarrow (2004), has persisted.[15] To test this effect over time, we construct separate yearly and monthly (using quarterly accounting data) observation interval data sets, following the procedures by Shumway (2001) and Chava and Jarrow (2004).

For the yearly observation interval, we merge data sets with annual firm-level accounting data from COMPUSTAT with monthly and daily equity price data from CRSP. A firm contributes an observation for every year after it starts appearing in the databases until the end of the sample period or its delisting year. To exploit data available in both CRSP and COMPUSTAT, we adjust our sample to contain 18,175 firms and 143,585 firm-year observations, which includes 3392 bankruptcies (according to our broad definition). In constructing the time-series variables, we lagged all independent variables one year, to ensure they were observable to the market at the beginning of each year. For example, if the dependent variable occurs in year t, we based the market information of independent variables on year $t - 1$, with the exception of SIGMA, which we computed as the sample standard deviation using the last 60 observable daily returns. The annual excess return (EXRET) is computed from the sum of monthly returns minus the value-weighted CRSP NYSE/AMEX/NASDAQ index return for the previous year. The relative size (RSIZ) is the logarithm of each firm's market equity value, divided by the weighted average of NYSE/AMEX/NASDAQ market equity value at the end of previous year. For the accounting variables (NI/TA, TL/TA, CASH/TA, and MB), we matched every yearly observation (calendar year) from the CRSP database with accounting data about the fiscal year, with a one-year lag. We removed any yearly observation with missing accounting variables. Similar to previous literature (Chava and Jarrow, 2004), we also discard outliers (e.g., typos, reporting errors) and truncate all variables at the 1st and

---

[15] Recall that their sample period is from 1963 to 1998.

99th percentiles of their pooled distributions across all firm-year observations. In addition, we replace any observation below the 1st percentile with the 1st percentile and any observation above the 99th percentile with the 99th percentile.

When dealing with the monthly observation interval, we still use daily and monthly equity price data from CRSP data, but instead of using the annual report in COMPUSTAT, we match the most recent previous quarterly accounting data to the current month. Thus both the accounting and market data are lagged by a month so that they are available to the market at the time of the estimation. The excess return (EXRET) is monthly returns minus the value-weighted CRSP NYSE/AMEX/NASDAQ index return at the end of the previous month; relative size (RSIZ) is the logarithm of each firm's market equity value divided by the weighted average of NYSE/AMEX/NASDAQ market equity value at the end of the previous month. SIGMA is the same as in the yearly data set. In line with Chava and Jarrow (2004), if the accounting data are not yet available, we use the previous quarter's accounting data.

However, unlike Chava and Jarrow (2004), we do not substitute a previously available observation for missing accounting or market data. When market or accounting data are missing for a given firm, it might indicate that the company in question prefers not to disclose the information or is unable to estimate certain items because of its poor financial situation, in which case the database institutions cannot report the information.[16] If we were to use previously available data to replace these missing values, it might bias our analysis; instead, we conservatively discard the observations with missing values[17] and thus ensure that the accounting data are

---

[16] In a credit card scoring context, Foster and Stine (2004) suggest that missing values are informative for assessing customer risks.

[17] For example, Lehman Brothers Holdings Inc. filed for Chapter 11 bankruptcy protection on September 15, 2008, the fourth quarter in that year. From our monthly observation sample set, in the third quarter of that year, two important accounting variables were missing: (1) cash and short-term

available at the beginning of the month. The final sample contains approximately 1.7 million firm-month observations and 2957 bankruptcies. The fewer bankruptcies in this observation data set, compared with the yearly set, results because we find more missing data in the quarterly accounting reports. To limit the influence of outliers even further, we winsorize all our explanatory variables using a 5th/95th percentile interval (as suggested by Campbell et al., 2008).

We report summary statistics for our 10 main explanatory variables for both firm-yearly and firm-monthly observations in Panels A and B of Table 2, respectively. The average value of EXRET is −7.33% per year (Panel A) and −1.36% per month (see Panel B), which reflects the underperformance of stocks during our sample period. Individual stock returns are extremely volatile; the average value of the annualized firm-level volatility (SIGMA) is 63% and 56% for the yearly and monthly data sets, respectively. The net income relative to adjusted total asset (NI/TA) is negative in both data sets, which indicates that firms had no profits during this period. Furthermore, the proportion of the total liability to adjusted total assets is approximately 50% for both data sets. That is, firms tend to match their liabilities with their equity.

[INSERT TABLE 2 HERE]

The summary statistics for bankrupt firms' yearly and monthly observations appear separately in Panels A and B of Table 3. This table indicates the characteristics of bankrupt firms one year and one month prior to their bankruptcy, which provides a hint about which interval predicts bankruptcy more accurately. Intuition tells us that if firms are about to file for bankruptcy, they should have low excess returns (EXRET), smaller relative firm size (RSIZ), larger volatility (SIGMA), low NI/TA, large TL/TA,

---

investments and (2) total liability. These missing data might have revealed the firm's financial distress, but databases could not record that information. We thus remove Lehman Brothers from our quarterly observation data set.

low CASH/TA, and large MB, compared with the summary statistics for all observations in Table 2.

[INSERT TABLE 3 HERE]

We find just these trends, for both yearly and monthly data. It is reasonable that firms about to go bankrupt have less net income, but they also might temporarily raise their net income reports by recognizing lower expenses to prevent investors from perceiving their financial distress. This strategy may bias bankruptcy predictions. In Table 3, the average NI/TA among bankrupt firm-year observations is –0.321, while the average for bankrupt firm-quarter observations, multiplied by 4, is around –0.232.[18] Firms appear to try to make their net income look better in the quarterly report, whereas the annual report provides more realistic accounting information. In turn, we expect that the data set with yearly observation intervals is more likely to provide greater explanatory power for bankruptcy models.

4. Bankruptcy Prediction Model

Following Shumway (2001), Chava and Jarrow (2004), and Campbell et al. (2008), we use a simple hazard model to estimate the probabilities of bankruptcy over the next period in a dynamic logit model. The main advantage of a simple hazard model is that we can control for and adjust the firm's at-risk period. For longer sampling periods, it is important to acknowledge that some firms file for bankruptcy after many years of being at risk, whereas other firms fail in their first risky year. Our model incorporates time-varying covariates that change over time; if a firm's financial health deteriorates before bankruptcy, its financial data reveal its changing health over time.

We assume that the marginal probability of bankruptcy in the next period follows

---

[18] The sum of four quarterly net income values equals annual net income, so to compare annual NI/TA and quarterly NI/TA, we must multiple quarterly NI/TA by 4 to approximate annual NI/TA.

a logistic distribution, expressed as:

$$P_{t-1}(Y_{it} = 1) = \frac{1}{1 + \exp(-\alpha - \beta x_{i,t-1})}, \quad (4)$$

with the parameters (α, β) and time-varying covariates $x_{i,t-1}$ for the discrete time hazard rate. The dependent variable Y_{it} equals 1 if the year observation with a delisting code belongs to our broad definition of bankruptcy, or 0 otherwise. A similar application for monthly data is Y_{it}, an indicator that equals 1 if the firms go bankrupt in month t. In addition, $x_{i,t-1}$ is a vector of explanatory variables known at the end of the previous month. To measure the explanatory power of a simple hazard model, we report McFadden's pseudo-R^2 coefficient, calculated as $1 - L_1/L_0$, where L_1 is the log-likelihood of the estimated model and L_0 is the log-likelihood of a null model that includes only a constant term. This hazard model is equivalent to a discrete time logit model with corrections in the estimation of the parameter's standard errors.

4.1. Bankruptcy Prediction with Industry Constraints

To provide empirical evidence regarding how industry constraints relate to bankruptcy, we estimate a dynamic logit model with five standard variables (EXRET, RSIZ, SIGMA, NI/TA, and TL/TA), two extra variables (CASH/TA and MB), and three industry constraint variables (H, $C_{i,B}$, and $C_{i,S}$) as regressors. We implement the dynamic logit model for both yearly and monthly observation intervals to learn whether a shorter interval data set can improve bankruptcy predictions. The yearly and monthly estimation results for firms traded on NYSE, AMEX, and NASDAQ during 1983–2008 appear in Panels A and B of Table 4.

[INSERT TABLE 4 HERE]

4.1.1. Yearly observation interval data set

In Panel A of Table 4, we note first that for market-driven variables (EXRET,

21

RSIZ, and SIGMA), and in agreement with prior findings, excess return and relative size relate negatively to bankruptcy predictions, whereas volatility has a positive impact. They are all statistically significant at the 1% level.[19] The signs of these coefficients are consistent with simple economic intuition; for example, smaller firms tend to use more short-term debt than larger firms, which make them riskier and more prone to financial distress and eventual bankruptcy. A firm with higher volatility bears more risk and thereby increases the possibilities of bankruptcy. For NI/TA and TL/TA though, our results are not consistent with Chava and Jarrow's (2004) claims that NI/TA is not significant when market-driven variables appear in the bankruptcy model. In regressions of our more recent data, accounting-driven variables relate significantly to bankruptcy, even if market variables already appear in the model. Both variables enter significantly and with the expected sign: Decreasing NI/TA (measure of firm profitability) and increasing TL/TA (measure of firm leverage) tend to increase chances of bankruptcy.

We next turn to the two added explanatory variables with accounting information. We observe CASH/TA is significantly and negatively related to bankruptcy across all models, which agrees with our expected intuition that a firm is more likely to go bankrupt when it cannot pay its creditors under liquidity limitations. Furthermore, MB meets our expectation; it shows a significantly positive sign in all models in Panel A of Table 4: The probability of bankruptcy increases when the market value is unusually high relative to the book value.

Table 4 also reveals how internal or external industry constraints can explain bankruptcies, beyond market-driven variables or accounting ratios. In Panel A, the Herfindahl index is significantly positive for bankruptcies in columns 1 and 4, which

---

[19] In our test, significance is less than 5% or 1% if the value of z-statistics is greater than 1.645 or 2.33, respectively.

is consistent with our first hypothesis, namely, that a higher concentration ratio (lower industry internal constraint) leads to a higher incidence of firm bankruptcy. Firms in highly concentrated industries are more likely to file for bankruptcy, even after we control for other market and accounting variables.

We also study the relationship between external industry constraints and bankruptcy prediction; according to Panel A of Table 4, in the first regression result (full model) and Model (3), the customer constraint ($C_{i,B}$) coefficient is positive and significant, in support of our second hypothesis. That is, the stronger customers' constraints, the higher the incidence of firm bankruptcy. Finally, the estimated coefficient for supplier constraints ($C_{i,S}$) is not significant in the full model but is negative and significant in Model (3), from which we removed the H variable. Thus the relationship between suppliers' constraints and bankruptcy remains unclear.

In summary, according to our yearly estimation results, we can not only reconfirm the effect of previously used bankruptcy variables but also provide empirical evidence in support of our two hypotheses.

4.1.2. Monthly observation interval data set

Following a similar procedure, we report the monthly estimation results in Panel B of Table 4. In general, these results are similar to our yearly outcomes, except that we find the CASH/TA variable is insignificant though still negative.

The value of the pseudo-R^2 based on yearly data also is 25% larger than that based on monthly data: 0.34 (last row, Panel A) compared with 0.27 (last row, Panel B), across all models in Table 4. Unlike Chava and Jarrow (2004), we find that the in-sample explanatory power of the bankruptcy prediction model based on yearly data is greater than that based on monthly data, perhaps because quarterly accounting reports do not deliver more accurate information pertaining to bankruptcy likelihood. Furthermore, more recent market information may incorrectly reflect firms' actual situation before

bankruptcies.

4.2. Robustness Tests: Yearly vs. Quarterly Data

4.2.1. In-sample performance

For a more robust comparison of in-sample performance, we recomputed the pseudo-R^2 for the years 1984–2008 by implementing the dynamic logit model with five standard variables (EXRET, RSIZ, SIGMA, NI/TA, and TL/TA). As Table 5 shows, the pseudo-R^2 from the yearly data is greater than that derived from monthly data for every year except 2004, 2006, and 2007. Both 2006 and 2007, just before the recent Great Recession, contain the lowest percentage of bankruptcies over our sample, according to Figure 1. Thus monthly data may perform slightly better than yearly data for predicting bankruptcies just before recessions. However, the advantage of yearly interval data remains pervasive, especially in 2003, one year after dot-com recession, and 1991–1994, covering two years of the 1990s recession and the next two years. We lack empirical results after 2008, but we maintain that yearly data are more useful than monthly data in normal times but especially in the years after recessions.

[INSERT TABLE 5 HERE]

Overall, the estimated coefficients are significant and with appropriate signs for both yearly and monthly data, except for CASH/TA, which becomes insignificant in the monthly data. The information from the past year seems more useful for predicting bankruptcies than information from the past month. Therefore, we use yearly data in our bankruptcy prediction model to determine how industry factors might explain bankruptcies.

4.2.2. Out-of-sample performance

To compare the out-of-sample performance of estimates from bankruptcy prediction models, we also analyze how well they forecast bankruptcy events before they appear, using only past available information. Specifically, we use the area under

the Receiver Operating Character (ROC) curve to assess prediction accuracy. This popular technique for validating internal credit rating models relies on the Area Under the ROC Curve (AUC) to indicate the quality of a rating model (Sobehart and Keenan, 2001). By construction, the model has no predictive power if the value of area under the ROC curve is 0.5 and demonstrates perfect forecasting ability if the value is 1.0. To compute the ROC measure, we estimate the model using 1983–1999 data and reserve the last nine years (2000–2008) as the out-of-sample period.[20] We use all available data before the out-of-sample period to estimate the parameters of each model, and then use the estimated parameters to forecast bankruptcy probabilities in 2000–2008 for each model. Finally, to calculate the AUC, we use forecasted bankruptcy probabilities and their corresponding observed status for each observation during the out-of-sample period.

Our aim is to investigate whether yearly interval data have better forecasting ability than monthly interval data, as we found for in-sample performance, so we implement the same model (logit regression, five variables). In addition, it is worthwhile to compare the yearly out-of-sample performance on market-driven (logit regression with EXRET, RSIZ, and SIGMA) and accounting-driven (logit regression with NI/TA and TL/TA) variable models against their monthly out-of-sample performance. The empirical AUC results appear in Table 6.

[INSERT TABLE 6 HERE]

In terms of the yearly interval, the AUC for the recent sample period with Shumway's model and the accounting-driven model is 0.885 and 0.79, respectively—greater than the AUC values reported by Chava and Jarrow (2004), that is, 0.868 and 0.723, respectively. The relatively larger difference in the

---

[20] Chava and Jarrow (2004) forecast bankruptcy probabilities of observations over an out-of-sample period of 1991–1999.

accounting-driven variable model may reflect two main influences. First, our sample is more recent. Second, we remove observations with missing previous year accounting data to avoid potential bias. Furthermore, the results in Table 6 indicate that the AUCs in yearly interval observations are greater than those in monthly data across all three models, consistent with our in-sample analysis. Therefore, we confirm that using yearly data is better than monthly data for both in-sample explanatory power and out-of-sample forecasting ability. Finally, we find that the AUC in the monthly, market-driven variable model (0.846) is higher than in Shumway's model (0.806). The performance differential reflects that addition of two more accounting variables; that is, additional quarterly accounting information cannot raise forecasting abilities and even might provide incorrect information. This result supports our previous supposition that earnings management efforts in quarterly accounting reports bias predictions of bankrupt events. In summary, with yearly data, the best model (in- and out-of-sample) includes both accounting and market-based variables. However, when we use monthly market data and quarterly accounting reports, the best model only features market data.

5. Bankruptcy Models Within and Across Industries

With industry-level measures (H, $C_{i,B}$, and $C_{i,S}$), we have connected bankruptcies to industry effects, including competition and bargaining power. In addition to firms' characteristics, greater internal concentration in an industry and less external bargaining power compared with other industries increase the possibilities of bankruptcy. To further this investigation, we next split the information related to our seven main firm-level variables (EXRET, RSIZ, SIGMA, NI/TA, TL/TA, CASH/TA, and MB) into within- and across-industry components. No previous research has outlined this relationship, though some studies tend to indicate that differences in firm characteristics are more important predictors of expected returns than are

industry-level variables. For example, Asness, Porter, and Stevens (2000) decompose book-to-market ratios into within- and across-industry components and find that the value effect is primarily within-industry, and Lewellen (1999) shows that a book-to-market factor explains common variation in returns unrelated to industry.

We adopt Asness et al.'s (2000) methodology to separate our market-wide independent variable[21] into two components: (1) the difference between firms' characteristics and the average characteristics of their industries (within-industry variables) and (2) the average characteristics of firms' industries (across-industry variables). This within- versus across-industry decomposition allows variables relative to their industry averages to reduce measurement error. For example, differences in accounting practices across industries can lead to differences in a variable that are unrelated to the incidence of bankruptcy.

The formal model is as follows: We regard X_{it} (market-wide variable) as any variable in our original logit model, but X_{Iit} (across-industry effect) denotes the average value of variable X over all firms located in a specific industry I. Then $(X_{it} - X_{Iit})$ is the deviation between the firm and its industry mean, or the within-industry effect. Our bankruptcy prediction model thus replaces any original explanatory variable by $(X_{it} - X_{Iit})$ and X_{Iit}, so our new prediction relies on within- and across-industry independent variables related to EXRET, RSIZ, SIGMA, NI/TA, TL/TA, CASH/TA, and MB.

5.1. *Bankruptcy Prediction Based on Within- and Across-Industry Variables*

Table 7 contains the logit model results for various specifications of the within- and across-industry variables. We consider three logit regressions for 1983–2008:

---

[21] Our market-wide independent variable includes three market-driven variables (EXRET, RSIZ, and SIGMA) and four accounting-driven variables (NI/TA, TL/TA, CASH/TA, and MB).

Column 1 estimates a regression on all within- and across-industry variables, Column 2 contains the regression for only within-industry variables, and Column 3 estimates the regression with only across-industry variables. The notations (W) and (A) in Table 7 refer to within- and across-industry variables, respectively.

[INSERT TABLE 7 HERE]

In Column 1, almost all within- and across-industry variables enter significantly and with a sign consistent with the market-wide variable in Panel A of Table 4. The first exception is that EXRET(A) and TL/TA(A) become insignificant, though still consistent with the expected sign. In contrast, MB(A) switches signs and is statistically significant. These results imply that the economic intuition perfectly fits the intra-industry scenario, but variation on industry-level characteristics sometimes cannot predict bankruptcy. In Column 2, all variables demonstrate the expected sign and are highly statistically significant, which reconfirms that a firm's characteristics, relative to its industry, are useful for gauging the incidence of bankruptcy. Not surprisingly, Column 3 shows that EXRET(A) becomes insignificant, and both TL/TA(A) and MB(A) become negatively significant, in contrast with our economic intuition. Moreover, we note that the value of pseudo-R^2 in the within-industry model is 0.33, but it is only 0.02 in the across-industry model. Thus it appears that bankruptcy prediction ability stems mainly from the variation of a firm's characteristics relative to its located industry, and industrial characteristics have very minor impact on bankruptcy predictions.

5.2. Out-of-Sample Performance

In Table 8 we compare AUC computed from predicted probabilities across all out-of-sample periods. According to Figure 1, there are five kinds of out-of-sample periods, based on three recessions (1990–1992, 2000–2002, and 2008) and two relatively stable periods between each pair of recessions (1993–1999 and 2003–2007).

For each out-of-sample period, we estimate each model with seven years of data before the predicted year.[22] The model notation is as follows: Model (1) is the logit regression on our main explanatory variables except for the industry constraint variables (EXERT, RSIZ, SIGMA, NI/TA, TL/TA, CASH/TA, and MB); W-and-A refers to the logit regression with all within- and across-industry variables; W is the logit regression with only within-industry variables; and A is the logit regression using only across-industry variables.

To compare forecasting abilities with respect to characteristics within and across industries, we also report the ratios of the AUC of the within-industry model to that of Model (1) (W/Model(1)) and of the AUC of the across-industry model to that of Model (1) (A/Model(1)). The AUC based on the W model is almost the same as that in Model (1), and the ratio is approximately 99% for all periods. This finding reinforces the extremely high explanatory power for the within-industry variables. Yet the ratio of AUC on A relative to Model (1) ranges from 59% to 69%; that is, across-industry effects provide a poor ability to predict the incidence of bankruptcy.

[INSERT TABLE 8 HERE]

6. Conclusion

We have examined two issues with this research: the effects of the degrees of competition and connection with customers and suppliers in a given industry on bankruptcy predictions and the relative performance of market-driven and accounting variables in terms of forecasting bankruptcy, as well as potential distortions in accounting information over time. Accordingly, we have proposed and tested two hypotheses. The first hypothesis states that greater concentration ratios lead to a higher incidence of bankruptcy for firms in that industry. We also predict that stronger

---

[22] We choose seven years as the in-sample length in accordance with our available in-sample data period for the first out-of-sample period, or 1983–1989.

customer constraints increase the incidence rates of bankruptcy, because these firms are more vulnerable to adverse economic conditions. We test and find support for both hypotheses. We also find that almost all the strength in bankruptcy predictions stems from the variation of the firm's characteristics relative to its industry, whereas industry-specific characteristics have a negligible impact on the accuracy of bankruptcy predictions.

Moreover, we find that models based on yearly data offer 25% more explanatory power, on average, than the models based on monthly data. When using yearly data, the best model (both in- and out-of-sample) integrates accounting and market-based variables; that is, both sources contain relevant information about bankruptcy likelihoods over time. However, when we use monthly market data and quarterly accounting reports, the best model only includes market data. These findings thus raise some challenging questions for ongoing research. In particular, the better forecasting performance of the accounting variables measured in yearly intervals, compared with quarterly intervals, may imply the potential existence of distortions in accounting signals at the quarterly level. The accounting variables might be distorted by shortcomings in the accounting conventions being applied, or because managers use excessive flexibility when reporting their accounting reports. But an alternative explanation is based on the relative efficiency of market prices in reflecting all publicly available information regarding bankruptcy because for monthly data, market variables are updated more quickly than are the quarterly accounting variables. For yearly data, updating of both market and accounting variables occurs at the same frequency. Further research should explore ways to disentangle the relative importance of the two explanations in yearly and quarterly data.

References

Asness, C., Porter, R.B., and Stevens, R., 2000. Predicting stock returns using industry relative firm characteristics. Working paper, University of Florida.

Bartov, E., Givoly, D., and Hayn, C., 2002. The rewards obtained when meeting the earnings expectations. *Journal of Accounting and Economics* 33, 173-204.

Bernhardt, D. and Campello, M., 2007. The dynamics of earnings forecast management. *Review of Finance* 11, 287-324.

Brander, J.A. and Lewis, T.R., 1986 Oligopoly and financial structure: The limited liability effect. *American Economic Review*, 76, 5, 956-970.

Brockman, P., and Turtle, H.J., 2003. A barrier option framework for corporate security valuation. *Journal of Financial Economics* 67, 511-529.

Burgstahler, D., and Dichev, I., 1997. Earnings management to avoid earnings decreases and losses. *Journal of Accounting and Economics* 24, 99-126.

Burt, R. S., 1983. *Corporate Profits and Cooptation*. New York: Academic Press.

Burt, R. S., 1988. The stability of American markets. *American Journal of Sociology* 94, 356-395.

Burt, R.S., Guilarte, M., Raider, H.J., and Yasuda, Y., 2002. Competition, contingency, and the external structure of markets. In P. Ingram & B. Silverman (Eds.). *New institutionalism in strategic management: Advances in strategic management* 19. New York: Elsevier, 167-217.

Campbell, J.Y., Hilscher, J., and Szilagyi, J., 2008. In search of distress risk. *Journal of Finance* 63, 2899-2939.

Chava, S., and Jarrow, R.A., 2004. Bankruptcy prediction with industry effects. *Review of Finance* 8, 537-569.

Cohen, R.B., Polk, C., and Vuolteenaho, T., 2003. The value spread**.** *Journal of Finance* 58. 609-642.

Davis, J.L., Eugene F.F., and Kenneth R.F., 2000. Characteristics, Covariances and Average Returns: 1929-1997. *Journal of Finance* 55, 389-406.

Dechow, P., Richardson, S., and Tuna, I., 2003. Why are earnings kinky? An examination of the earnings management explanation. *Review of Accounting Studies* 8, 355-384.

Degeorge, F., Patel , J., and Zeckhauser, R., 1999. Earnings manipulations to exceed thresholds. *Journal of Business* 72, 1-33.

Dichev, I., 1998. Is the risk of bankruptcy a systematic risk? *Journal of Finance* 53, 1131-1147.

Foster, D.P. and Stine, R.A., 2004. Variable selection in data mining: Building a predictive model for bankruptcy. *Journal of the American Statistical Association* 99, 303-313.

Hayn, C., 1995. The information content of losses. *Journal of Accounting and Economics* 20, 125-153.

Healy, P.M. and Wahlen, J.M., 1999. A review of the earnings management literature and its implications for standard setting. *Accounting Horizons* 13, 365–383.

Hertzel, M.G., Li, Z., Officer, M.S., and Rodgers, K.J., 2008. Inter-firm linkages and the wealth effects of financial distress along the supply chain. *Journal of Financial Economics* 87, 374-387.

Hou, K. and Robinson, D.T., 2006. Industry concentration and average stock returns. *Journal of Finance* 61, 1927–1956.

Kale, J.R. and Shahrur, H., 2007. Corporate capital structure and the characteristics of suppliers and customers. *Journal of Financial Economics* 83, 321-365.

Kinney, W.R., Jr. and McDaniel, L.S., 1989. Characteristics of firms correcting previously reported quarterly earnings. *Journal of Accounting and Economics* 11, 71–93.

Laitinen, E.K. and Laitinen, T., 1998. Cash management behaviour and failure prediction. *Journal of Business Finance and Accounting* 25, 893-919.

Lewellen, J., 1999. The time-series relations among expected return, risk, and book-to-market. *Journal of Financial Economics* 54, 5-43.

Maksimovic, V. and Phillips, G., 1998. Asset efficiency and reallocation decisions of bankrupt firms. *Journal of Finance* 53, 1495–1532.

Matsunaga, S.R. and Park, C.W., 2001. The effect of missing a quarterly earnings benchmark on the CEO's annual bonus. *The Accounting Review* 76, 313–332.

Myers, J., Myers, L., and Skinner, D., 2007. Earnings momentum and earnings management. *Journal of Accounting, Auditing and Finance* 22, 249–284.

Opler, T.C., and Titman, S., 1994. Financial distress and corporate performance. *Journal of Finance* 49, 1015-1040.

Roychowdhury, S., 2006. Earnings management through real activities manipulation. *Journal of Accounting and Economics* 42, 335-370.

Sarig, O., 1998. The effect of leverage on bargaining with a corporation. *Financial Review* 33, 1–16.

Shih, H.-Y., 2007. Structural analysis of market structure and performance: A new empirical industrial structure method in competitive analysis. International Conference on Business and Information, Tokyo, Japan.

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

Sobehart, J., and Keenan, S., 2001. Measuring default accurately. *Risk*, March, 31-33.

Talmud, I., 1994. Relations and profits: The social organization of Israeli industrial competition. *Social Science Research* 23, 109-135.

Titman, S. and Wessels, R., 1988. The determinants of capital structure choice. *Journal of Finance* 43, 1-19.

Yasuda, Y., 2005. *Social organization and performance inequality in Japanese and American markets*. Tokyo: Manufacturing Management Research Center.

BANKRUPTCIES AS PERCENTAGE OF ACTIVE COMPANIES
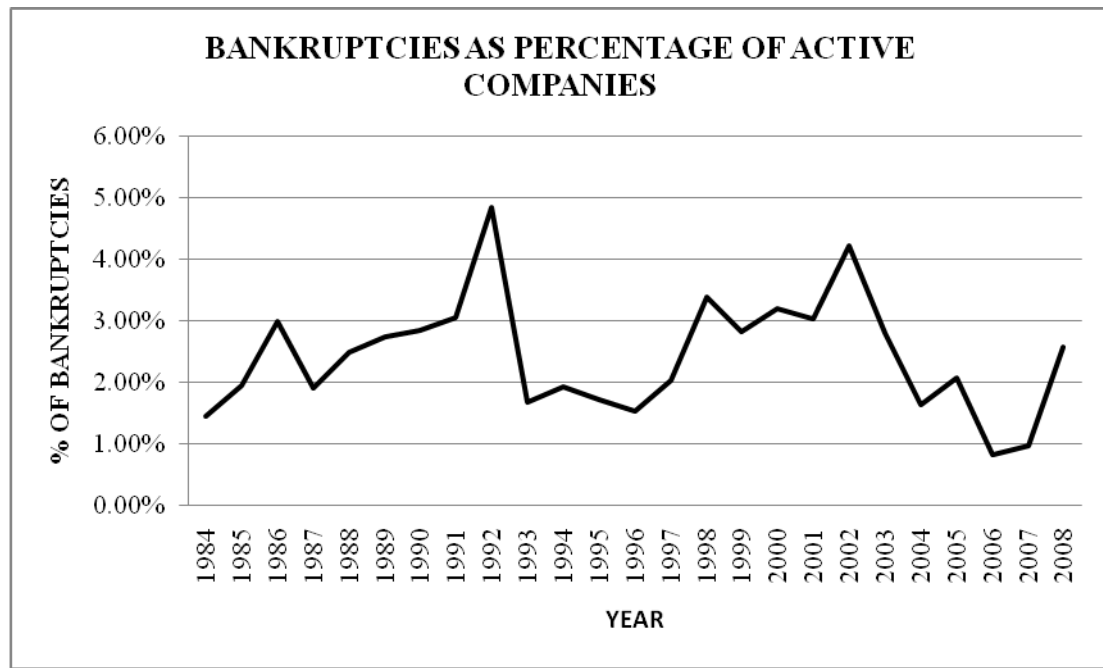
YEAR	# of bankrupt firms	# of active firms	% of bankrupt to active firms	YEAR	# of bankrupt firms	# of active firms	% of bankrupt to active firms
1984	61	4233	1.44%	1997	148	7349	2.01%
1985	84	4327	1.94%	1998	246	7279	3.38%
1986	135	4530	2.98%	1999	194	6889	2.82%
1987	90	4742	1.90%	2000	220	6922	3.18%
1988	122	4919	2.48%	2001	180	5962	3.02%
1989	134	4922	2.72%	2002	253	6023	4.20%
1990	138	4870	2.83%	2003	153	5475	2.79%
1991	140	4602	3.04%	2004	91	5602	1.62%
1992	237	4894	4.84%	2005	113	5530	2.04%
1993	87	5226	1.66%	2006	45	5509	0.82%
1994	121	6321	1.91%	2007	52	5442	0.96%
1995	113	6641	1.70%	2008	130	5072	2.56%
1996	105	6912	1.52%				

Figure 1. Bankruptcy by Year. Bankrupt firms must satisfy two conditions: (1) Firms have been listed publicly in the NYSE, AMEX, or NASDAQ exchange during 1984–2008 and (2) their reported delisting code in CRSP is 400 or 550–585. The 3392 bankruptcies in our sample are separated by year in this table, along with bankruptcies as a percentages of active number of firms in that year.

BANKRUPTCY BY EXCHANGE

NASDAQ   2723 (80.28%)
AMEX   363 (10.70%)
NYSE   306 (9.02%)
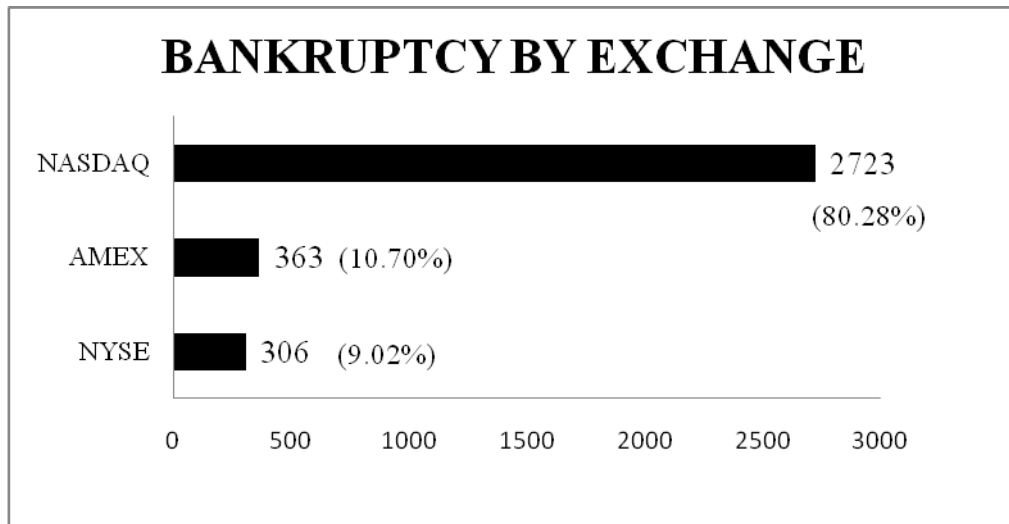
Figure 2. Bankruptcy by Exchange. Bankrupt firms must satisfy two conditions: (1) Firms have been listed publicly in the NYSE, AMEX, or NASDAQ exchange during 1984–2008 and (2) their reported delisting code in CRSP is 400 or 550–585. The 3392 bankruptcies in our sample are separated by exchange in this table, along with the percentages of the total number of bankruptcies in brackets.

Table 1. Bankruptcy by IO Code. Bankrupt firms must satisfy three conditions: (1) Firms have been listed publicly in the NYSE, AMEX, or NASDAQ exchange during 1984 to 2008; (2) their reported delisting code in CRSP is 400 or 550–585; and (3) this SIC code mirrors an IO code in the conversion table. The 3392 bankruptcies in our sample are listed by IO code, along with the percentages of the total number of bankruptcies. The industry classification system changed after 1998, so there are only eight industries before 1997 and fourteen industries after 1998. Panel A reports the number of bankruptcies from 1984 to 1997, and Panel B refers to 1998–2008.

	Panel A: Bankruptcies from 1984 to 1997		
IOCODE	INDUSTRY NAME	# OF BANKRUPT FIRMS	% OF BANKRUPT FIRMS
1	Agriculture, forestry, fishing and hunting	19	1.11%
2	Mining	197	11.49%
3	Construction	49	2.86%
4	Manufacturing	645	37.61%
5	Transportation, communication and utilities	82	4.78%
6	Trade	207	12.07%
7	Finance, insurance, real estate, rental, and leasing	190	11.08%
8	Services	326	19.01%
	Total	1715	100.00%
	Panel B: Bankruptcies from 1998 to 2008		
1	Agriculture, forestry, fishing and hunting	5	0.30%
2	Mining	63	3.76%
3	Utilities	9	0.54%
4	Construction	43	2.56%
5	Manufacturing	665	39.65%
6	Wholesale trade	77	4.59%

7	Retail trade	69	4.11%
8	Transportation and warehousing	45	2.68%
9	Information	253	15.09%
10	Finance, insurance, real estate, rental, and leasing	204	12.16%
11	Professional and business services	141	8.41%
12	Educational services, health care, and social assistance	27	1.61%
13	Arts, entertainment, recreation, accommodation, and food services	67	4.00%
14	Other services, except government	9	0.54%
	Total	1677	100.00%

Table 2. Summary Statistics: All Firms. This table offers summary statistics (annual and monthly) of the independent variables for the sample period 1983–2008. The sample has relevant information in COMPUSTAT and CRSP and firms listed on all the three main exchanges (NYSE, AMEX, and NASDAQ) are considered. There are 18,175 firms and 143,585 observations in the yearly sample and 18,951 firms and 1,761,849 observations in the monthly sample. Each firm year and firm month is considered a separate observation for the yearly and monthly samples. The variable definitions are: (1) EXRET = excess annual return over the value-weighted NYSE, AMEX and NASDAQ return; (2) RSIZ = relative size, measured as the log of the firm's market capitalization divided by total NYSE, AMEX, NASDAQ market cap; (3) SIGMA = standard deviation computed by using the last sixty observable daily returns; (4) NI/TA = net income/adjusted total assets; (5) TL/TA = total liabilities/adjusted total assets; (6) CASH/TA = ratio of a company's cash and short-term assets to its adjusted total assets; (7) MB = market-to-book ratio; (8) H = Herfindahl index; (9) $C_{i,B}$ = buyer constraint; and (10) $C_{i,S}$ = supplier constraint. Panels A and B refer to all firm-year observations and all firm-month observations, respectively.

Variable	Median	Mean	Min	Max	Std
Panel A: Firm-year observations, 1983 to 2008					
EXRET	-0.0085	-0.0733	-2.2218	1.0297	0.5164
RSIZ	-11.0844	-10.9744	-15.2721	-5.9389	2.0269
SIGMA	0.4793	0.6393	0.0827	2.8983	0.5128
NI/TA	0.0226	-0.0206	-0.9704	0.2348	0.1802
TL/TA	0.5118	0.5123	0.0261	1.1543	0.268
CASH/TA	0.0690	0.1415	0.0001	0.7660	0.1720
MB	1.6072	3.0628	0.0287	45.6008	5.4033
H	0.0139	0.0229	0.0079	0.2996	0.0238
$C_{i,B}$	0.0192	0.0190	0.0095	0.0454	0.0043
$C_{i,S}$	0.0202	0.0204	0.0079	0.0356	0.0059
Panel B: Firm-month observations, 1983 to 2008					
EXRET	-0.0109	-0.0136	-0.2653	0.2237	0.1199
RSIZ	-11.0928	-10.9779	-14.1191	-7.3451	1.8875
SIGMA	0.4592	0.5607	0.1657	1.4332	0.3476
NI/TA	0.0054	-0.0040	-0.1175	0.0399	0.038
TL/TA	0.5066	0.5052	0.086	0.9299	0.2583
CASH/TA	0.0633	0.1334	0.0025	0.5449	0.1567
MB	1.6432	2.5105	0.3756	9.7558	2.3811
H	0.0161	0.0281	0.0090	0.4208	0.0314
$C_{i,B}$	0.0201	0.0220	0.0089	0.0771	0.0082
$C_{i,S}$	0.0238	0.0239	0.0077	0.0548	0.0091

Table 3. Summary Statistics of Bankrupt Firms. This table offers summary statistics (annual and monthly) of the independent variables for the sample period 1983–2008. There are 3392 and 2957 bankruptcies with relevant information in COMPUSTAT and CRSP for the yearly and monthly data sets, respectively, and firms listed on all the three main exchanges (NYSE, AMEX, and NASDAQ) are considered. Each firm year is considered a separate observation. The variable definitions are: (1) EXRET = excess annual return over the value-weighted NYSE, AMEX and NASDAQ return; (2) RSIZ = relative size, measured as the log of the firm's market capitalization divided by total NYSE, AMEX, NASDAQ market cap; (3) SIGMA = standard deviation computed by using the last sixty observable daily returns; (4) NI/TA = net income/adjusted total assets; (5) TL/TA = total liabilities/adjusted total assets; (6) CASH/TA = ratio of a company's cash and short-term assets to its adjusted total assets; and (7) MB = market-to-book ratio. Panels A and B refer to bankrupt firm-year observations and bankrupt firm-month observations, respectively.

Variable	EXRET	RSIZ	SIGMA	NI/TA	TL/TA	CASH/TA	MB
Panel A: Bankrupt firm-year observations, 1983 to 2008							
Median	-0.288	-13.653	1.447	-0.202	0.682	0.055	1.466
Mean	-0.434	-13.505	1.515	-0.321	0.658	0.130	5.130
Std	0.815	1.373	0.86	0.348	0.312	0.175	9.537
Panel B: Bankrupt firm-month observations, 1983 to 2008							
Median	-0.112	-14.119	1.433	-0.057	0.779	0.042	1.015
Mean	-0.089	-13.695	1.159	-0.058	0.686	0.113	2.457
Std	0.172	0.798	0.403	0.053	0.267	0.154	3.001

Table 4. Logit Regressions of Bankruptcy Indicators on Predictor Variables. For the hazard model proposed by Shumway (2001), the estimates include firms traded on the NYSE, AMEX, and NASDAQ with yearly observations from 1983 to 2008, a total of 143,585 yearly observations of 18,175 firms. The model includes five explanatory variables suggested by Shumway (2001): EXRET, RSIZ, SIGMA, NI/TA, and TL/TA. In addition, it features CASH/TA and MB, as well as three independent variables for industry constraints (H, $C_{i,B}$, and $C_{i,S}$). Parameter estimates are given first, followed by the absolute value of z-statistics reported in parentheses. *Significant at 5%. **Significant at 1%. The McFadden R-square is in the final row.

	Full Model	Shumway's Model	Model (1)	Model (2)	Model (3)
Panel A: All firm-year observations, 1983 to 2008					
Intercept	-14.32	-13.77	-14.01	-14.13	-14.14
	(63.49)**	(70.05)**	(67.33)**	(67.71)**	(62.98)**
EXRET	-0.18	-0.15	-0.18	-0.18	-0.17
	(6.73)**	(5.72)**	(6.79)**	(6.77)**	(6.57)**
RSIZ	-0.65	-0.63	-0.65	-0.65	-0.66
	(42.48)**	(42.28)**	(42.67)**	(42.40)**	(42.56)**
SIGMA	0.92	0.94	0.92	0.92	0.92
	(36.16)**	(36.96)**	(36.39)**	(36.13)**	(36.28)**
NI/TA	-1.95	-2.08	-1.95	-1.95	-1.95
	(29.13)**	(34.38)**	(29.06)**	(29.09)**	(29.06)**
TL/TA	1.74	1.87	1.70	1.74	1.72
	(23.29)**	(27.97)**	(22.99)**	(23.47)**	(23.02)**
CASH/TA	-0.53		-0.63	-0.54	-0.59
	(4.13)**		(4.92)**	(4.26)**	(4.56)**
MB	0.03		0.03	0.03	0.03
	(11.39)**		(11.08)**	(11.34)**	(11.04)**
H	5.49			5.56	
	(8.08)**			(8.51)**	
$C_{i,B}$	13.32				18.07
	(2.44)**				(3.27)**
$C_{i,S}$	-4.04				-12.68
	(0.98)				(3.12)**
Pseudo-R^2	0.344	0.338	0.342	0.344	0.342

Panel B: All firm-month observations, 1983 to 2008

	Full Model	Shumway's Model	Model (1)	Model (2)	Model (3)
Intercept	-25.35	-25.08	-25.26	-25.29	-25.32
	(62.14)**	(64.23)**	(62.58)**	(62.67)**	(62.06)**
EXRET	-0.98	-0.96	-0.98	-0.98	-0.98
	(8.32)**	(8.19)**	(8.36)**	(8.35)**	(8.32)**
RSIZ	-1.24	-1.23	-1.24	-1.24	-1.24
	(40.88)**	(41.19)**	(40.86)**	(40.81)**	(40.81)**
SIGMA	1.44	1.44	1.44	1.44	1.44
	(24.72)**	(24.79)**	(24.81)**	(24.74)**	(24.79)**
NI/TA	-8.31	-8.53	-8.31	-8.31	-8.31
	(20.89)**	(22.83)**	(20.88)**	(20.89)**	(20.88)**
TL/TA	2.41	2.46	2.40	2.41	2.41
	(28.50)**	(32.12)**	(28.45)**	(28.54)**	(28.40)**
CASH/TA	-0.13		-0.19	-0.16	-0.16
	(0.95)		(1.36)	(1.13)	(1.11)
MB	0.02		0.02	0.02	0.02
	(3.02)**		(3.03)**	(3.06)**	(2.97)**
H	1.50			1.65	
	(2.85)**			(3.17)**	
$C_{i,B}$	6.53				7.69
	(2.06)*				(2.45)**
$C_{i,S}$	-4.93				-5.93
	(1.76)				(2.12)*
Pseudo-R^2	0.278	0.277	0.278	0.278	0.278

Table 5. Comparison of Pseudo-R^2 in Yearly and Monthly Data. This table reports McFadden's pseudo-R^2 obtained from the logit model with five standard variables (EXRET, RSIZ, SIGMA, NI/TA, and TL/TA) for each year through 1984 to 2008. McFadden's pseudo-R^2 coefficient is calculated as $1 - L_1/L_0$, where L_1 is the log-likelihood of the estimated model and L_0 is the log-likelihood of a null model that includes only a constant term.

YEAR	Yearly-R^2	Monthly-R^2	YEAR	Yearly-R^2	Monthly-R^2
1984	0.3272	0.2452	1997	0.4466	0.3213
1985	0.3277	0.2672	1998	0.4078	0.3297
1986	0.4106	0.3232	1999	0.3408	0.3192
1987	0.3363	0.3183	2000	0.3237	0.2925
1988	0.3090	0.2399	2001	0.3262	0.2712
1989	0.2507	0.2413	2002	0.3527	0.2960
1990	0.3737	0.2803	2003	0.4456	0.2852
1991	0.3644	0.2356	2004	0.3042	0.3130
1992	0.4032	0.2896	2005	0.3212	0.2459
1993	0.4028	0.2931	2006	0.1649	0.1956
1994	0.4193	0.3076	2007	0.2574	0.3571
1995	0.4121	0.3060	2008	0.2800	0.2731
1996	0.4321	0.3294			

Table 6. Forecasting Accuracy, 2000–2008, by AUC. The dynamic logit model is estimated with data from 1983–1999 with yearly and monthly observation intervals to forecast bankruptcy probabilities for 2000–2008. This table reports the area under the ROC curve (AUC) obtained from the Shumway's model, the market-driven variable model, and the accounting-driven variable model. The AUC is calculated using forecasted bankruptcy probabilities and their corresponding observed status for each observation over the out-of-sample period. We provide AUC results based on yearly and monthly interval observation in the first and second rows, respectively.

	Shumway's model	Market-driven model	Accounting-driven model
		AUC	
Yearly	0.885	0.878	0.790
Monthly	0.806	0.846	0.693

Table 7. Bankruptcy Indicators, Within- and Across-Industry Variables. The hazard model proposed by Shumway (2001) is used for within- and across-industry variables. The model is estimated for firms traded on the NYSE, AMEX, and NASDAQ with yearly observations from 1983 to 2008, for a total of 143,585 yearly observations of 18,175 firms. The model includes five explanatory variables suggested by Shumway (2001), EXRET, RSIZ, SIGMA, NI/TA, and TL/TA, as well as CASH/TA, MB, and the three industry constraint variables: H, $C_{i,B}$, and $C_{i,S}$. Except for industry constraint variables, each explanatory variable (X_i) is decomposed into a within-industry variable, denoted as $(X_i - X_{Ii})$, and an across-industry variable (X_I). The notations W and A stand for within-industry and across-industry variables, respectively. We consider three logit regressions. (1) Within and Across: the regression includes all within- and across-industry variables; (2) Within: the regression includes only within-industry variables; and (3) Across: the regression includes only across-industry variables. Parameter estimates are given first, followed by the absolute value of the z-statistics reported in parentheses. *Significant at 5%. **Significant at 1%. The McFadden R-square is in the final row.

Variable	Within and Across	Within	Across
Intercept	-12.26	-5.20	-8.92
	(20.64)**	(132.31)**	(17.02)**
EXRET (W)	-0.17	-0.16	
	(6.27)**	(5.98)**	
EXRET (A)	-0.24		-0.01
	(1.84)		(0.11)
RSIZ (W)	-0.65	-0.61	
	(41.31)**	(40.19)**	
RSIZ (A)	-0.63		-0.51
	(12.80)**		(11.54)**
SIGMA (W)	0.95	1.01	
	(35.00)**	(37.42)**	
SIGMA (A)	0.66		0.79
	(5.67)**		(7.91)**
NI/TA (W)	-1.88	-1.95	
	(26.88)**	(28.31)**	
NI/TA (A)	-4.02		-5.27
	(5.88)**		(9.16)**
TL/TA (W)	1.79	1.81	
	(23.19)**	(23.54)**	
TL/TA (A)	0.19		-0.88
	(0.71)		(3.81)**
CASH/TA (W)	-0.47	-0.54	
	(3.59)**	(4.15)**	
CASH/TA (A)	-2.67		-2.75
	(3.96)**		(4.61)**

MB (W)	0.03	0.02	
	(11.42)**	(9.87)**	
MB (A)	-0.08		-0.08
	(2.60)**		(2.85)**
Pseudo-R^2	0.34	0.33	0.02

Table 8. Out-of-Sample Performance, Within- and Across-Industry Predictors. This table compares out-of-sample performance to within- and across-industry predictors by measuring forecasting accuracy. An ROC curve is constructed from the predicted probabilities computed for the out-of-sample period. The five kinds of out-of-sample periods reflect three recessions (1990–1992, 2000–2002, and 2008) and two relatively stable periods (1993–1999 and 2003–2007). For each out-of-sample period, each model is estimated with seven years of data before the predicted year. Model(1) is the logit regression on our main explanatory variables except for industry constraint variables (EXERT, RSIZ, SIGMA, NI/TA, TL/TA, CASH/TA, and MB); W-and-A is the logit regression on all within- and across-industry variables; W is the logit regression on only within-industry variables; and A is the logit regression on only across-industry variables. Moreover, to compare the forecasting abilities with respect to factors within and across industries, we report the ratio of the AUCs in the within-industry model to Model(1) (W/Model(1)) and in the across-industry model to Model(1) (A/Model(1)).

	Model(1)	W-and-A	W	A	W/Model(1)	A/Model(1)
Out-of-Sample Period				AUC		
1990-1992	0.934	0.933	0.923	0.585	98.82%	62.63%
1993-1999	0.919	0.919	0.908	0.546	98.80%	59.41%
2000-2002	0.891	0.896	0.887	0.600	99.55%	67.34%
2003-2007	0.887	0.848	0.880	0.619	99.21%	69.79%
2008	0.884	0.889	0.872	0.552	98.64%	62.44%

Appendix
Delisting codes in the CRSP Database

Code	Definition
400	Issue stopped trading as result of company liquidation
550	Delisted by current exchange - insufficient number of market makers
551	Delisted by current exchange - insufficient number of shareholders
552	Delisted by current exchange - price fell below acceptable level
560	Delisted by current exchange - insufficient capital, surplus, and/or equity.
561	Delisted by current exchange - insufficient (or non-compliance with rules of) float or assets.
570	Delisted by current exchange - company request (no reason given).
572	Delisted by current exchange - company request, liquidation.
573	Delisted by current exchange - company request, deregistration (gone private).
574	Delisted by current exchange - bankruptcy, declared insolvent
575	Delisted by current exchange - company request, offer rescinded, issue withdrawn by underwriter
580	Delisted by current exchange - delinquent in filing, non-payment of fees
581	Delisted by current exchange - failure to register under 12G of Securities Exchange Act
582	Delisted by current exchange - failure to meet exception or equity requirements.
583	Delisted by current exchange - denied temporary exception requirement.
584	Delisted by current exchange - does not meet exchange's financial guidelines for continued listing.