INDUSTRY EFFECTS ON STOCK PRICE CRASH RISK

QUANTIFYING INDUSTRY EFFECTS AND THE EFFECT OF STRUCTURAL INDUSTRY CHARACTERISTICS

INDRA BERLIN

TUVA BROSTRÖM

Bachelor Thesis

Stockholm School of Economics

2023

Industry Effects on Stock Price Crash Risk: Quantifying Industry Effects and the Effect of Structural Industry Characteristics

Abstract:

This paper empirically investigates the determinants of stock price crash risk at the industry level. Prior research has largely focused on firm-level and macro-level factors that impact stock price crash risk, with little attention being drawn to inter-industry differences. Using a large sample of US firms from different industries, this paper finds significant industry fixed effects on stock price crash risk, with varying magnitudes between industries. For example, industries in the health care sector are consistently associated with high crash risk, while industries in the energy sector are consistently associated with low crash risk. Furthermore, this study examines possible underlying dynamics of these industry effects by analyzing the effects of three structural industry characteristics on crash risk. The findings indicate a significant positive effect of industry growth rate on stock price crash risk, as well as a significant negative effect of industry profitability during the period 1985-2000. No effects are found for industry concentration, although it is concluded that this may be due to limitations in the proxy used.

Keywords:

Stock Price Crash Risk, Industry Level Effects, Industry Concentration, Industry Growth Rate, Industry Profitability

Authors:

Indra Berlin (25006)
Tuva Broström (25028)

Tutors:

Riccardo Sabbatucci, Assistant Professor, Department of Finance

Examiner:

Adrien d’Avernas, Assistant Professor, Department of Finance

Bachelor Thesis

Bachelor Program in Business and Economics
Stockholm School of Economics

1. Introduction

In recent years, stock price crash risk has been a topic of interest to researchers, and a significant amount of research has been devoted to understanding its determinants. Interest in such research has been further amplified by high-profile corporate scandals that have resulted in stock crashes, such as WorldCom, Enron, and Xerox (Chang et al., 2017). Investors demand higher expected returns for stocks with more negative skewness, which implies crash risk is a priced risk factor (Harvey and Siddique, 2000; Conrad et al., 2013). Hence, understanding the causes of stock price crash risk has important implications for asset pricing.

A large part of the recent literature on stock price crash risk has been based on Jin and Meyers' (2006) agency theory framework, which outlines managers' accumulation of bad news as an important cause of stock price crashes. According to this framework, managers are incentivized to withhold bad news from the market until it is no longer possible to do so anymore. This results in the bad news being released all at once, which in turn causes stock prices to decline significantly.

Previous research on the determinants of stock price crash risk provides empirical support for the agency theory framework. Studies have mainly focused on firm-specific factors that alter managers' incentives in ways that lead to bad news hoarding and consequently increases future crash risk. These factors include firm size, financial statement transparency, tax avoidance, CFO equity incentives, and stock liquidity (Chen et al., 2001; Hutton et al., 2009; Kim et al., 2011a; Kim et al., 2011b; and Chang et al., 2017)). Additionally, macro-level factors have also received considerable attention, such as country-level religiosity and environmental monitoring (Callen and Fang, 2015; and Zhang et al., 2021). However, there has been little research focusing on inter-industry differences and the impact of industry-level characteristics on stock price crash risk. Although previous research has included controls for industry effects (e.g. Chang et al., 2017; Hutton et al., 2009), the significance of these effects has not been drawn to attention. This represents a gap in the literature which this research paper aims to address.

Understanding inter-industry differences in stock price crash risk and the underlying industry characteristics that determine these differences is of interest to investors, managers and policymakers who seek to mitigate the potential damage that stock price crashes inflict upon shareholder wealth. Additionally, since crash risk is a priced risk factor, understanding its determinants has implications for empirical asset pricing.

There are several industry-level factors that may affect stock price crash risk. Broadly, these factors can be divided into operational industry characteristics and structural industry characteristics, and in this paper we focus our analysis on the latter. We choose three structural industry characteristics that could potentially be relevant for stock price crash risk: (1) Industry concentration, (2) Industry growth, and (3) Industry profitability.

Market concentration serves as a commonly used proxy for product market competition. Previous research has found that competitive pressure is associated with higher crash risk (Li and Zang, 2019). Based on Jin and Meyers' (2006) agency theory approach to stock price crash risk, firms operating in industries with high degrees of product market competition could be expected to have a higher stock price crash risk, as competition incentivizes managers to delay bad news releases.

Another structural industry-level factor to consider is industry growth rate. A high growth rate is considered an indicator of industry disequilibrium (Yip, 1982), and firms operating in growth industries may experience high levels of turbulence (Agarwal & Gort, 1996). This implies that the consequent uncertainty and rapid change in market conditions in high-growth industries could lead to an increased risk of stock price crashes.

Industry profitability is also worth examining as a possible determinant of stock price crash risk. Managers of firms in industries with low profitability could have an incentive to withhold bad news from the market, due to a multitude of strong competitive forces these firms are faced with (Porter, 2008). Hence, low industry profitability could be expected to increase stock price crash risk.

This paper aims to examine industry-level effects on stock price crash risk by analyzing a large sample of U.S. firms in different industries over the period 1985-2022. To ensure the robustness of our findings, we measure crash risk using three different variables: one variable that captures extremely low negative skewness in a stock's return distribution, one dummy variable indicating if a stock experiences at least one week with extremely low returns during each fiscal year, and one variable capturing extreme down-to-up volatility.

The empirical investigation is divided into two parts. First, we assess whether there are significant industry effects on stock price crash risk. To achieve this, we conduct regressions that control for industry fixed effects, allowing us to examine the significance and magnitude of these effects on crash risk. Our results indicate significant industry fixed effects on crash risk for a majority of industries, with varying magnitudes. These findings suggest that different industries exhibit different levels of stock price crash risk, even when accounting for firm-specific characteristics known to be correlated with crash risk. For instance, industries in the health care sector consistently display relatively high crash risk, while industries in the energy sector exhibit relatively low crash risk. The financial sector and consumer discretionary sector show associations with both relatively high and low crash risk. We also identify industries with significant results of the highest and lowest magnitudes, and plot crash risk for those industries over time. Through this univariate analysis, we observe that the established industry effect appears to differ between the periods 1985-2000 and 2001-2022.

In the second part of the empirical analysis, we explore potential underlying industry-level factors that may explain the observed industry fixed effects. In this paper, we focus on structural industry characteristics rather than operational characteristics. Specifically, we examine industry concentration, industry growth rate, and industry profitability as industry-specific characteristics. We conduct several regressions with these characteristics included as explanatory variables.

Our findings indicate that industry concentration has no significant effect on crash risk. However, when conducting robustness tests using an alternative industry concentration measure, we observe significant negative effects. This suggests that the insignificant results may be attributed to our choice of variable, necessitating further research to draw definitive conclusions regarding the relationship between industry concentration and crash risk. Additionally, we find a significant positive effect of industry growth rate on crash risk, which remains robust across different measures of crash risk. This suggests that stocks in an industry with higher industry growth rates are associated with elevated crash risk. However, this effect appears to be driven by the period 2001-2022. Furthermore, we find no effect of industry profitability on crash risk for the entire sample period, but a significant negative effect during the period 1985-2000. This indicates that high industry profitability during that period increased crash risk. Given the limited research on industry-level effects on stock price crash risk, the existing literature neither confirms nor contradicts our findings. However, industry concentration has been found to have a negative effect on crash risk in a study of the Chinese market (Li and Luo, 2020), as well as a positive effect in the U.S. market during the period 1998-2009 (Li and Zang, 2019), which contradicts our insignificant results in the U.S. market.

The remaining sections of this paper are organized as follows. Section 2 provides a brief review of the related literature. Section 3 presents our hypotheses. Section 4 discusses the sample data and describes how all variables were constructed. Section 5 presents our empirical analysis and provides interpretations of the regression results, centered around a plurality of multivariate regression analyses examining the impact of sector and industry classifications on crash risk, as well as regressions exploring the relationship between selected industry characteristics and crash risk. Finally, section 6 concludes.

2. Related Literature

The majority of the existing literature on the subject of stock price crash risk adheres to Chen et al.'s (2001) definition, which characterizes crash risk as "the conditional skewness of the return distribution".[1] Subsequent studies have primarily focused on identifying the determinants of crash risk, adopting an agency theory framework as outlined by Jin and Myers (2006). According to their perspective, a notable predictor of stock price crash risk is the practice of concealing unfavorable news by firm insiders, which arises due to information asymmetries and conflicting interests. Such concealment may be motivated by factors such as corporate tax avoidance and equity-based incentives (Kim et al., 2011a; Kim et al., 2011b).

Managers often tend to promptly disclose positive news to investors while delaying the release of negative news. This behavior is exacerbated when managers face heightened career concerns and possess a substantial personal stake in the company (Kothari et al., 2009). Managers may defer the disclosure of adverse news with the hope that future positive developments or improved performance could mitigate the impact of undisclosed negative information (Graham et al., 2005; Kothari et al., 2009). Consequently, the accumulation of undisclosed negative news can result in a stock price crash upon its eventual revelation in the market. Empirical research has lent support to this theoretical framework, exploring various factors that influence managers' propensity to conceal negative news and, by extension, their association with crash risk. Among the firm-level factors identified as having an impact on crash risk are stock price volatility, past stock returns, firm size, financial statement transparency, tax avoidance, CFO equity incentives, and stock liquidity (Chen et al., 2001; Hutton et al., 2009; Kim et al., 2011a; Kim et al., 2011b; and Chang et al., 2017)). Additionally, certain macro-level factors, such as country-level religiosity and environmental monitoring, have been found to be relevant in understanding crash risk (Callen and Fang, 2015; and Zhang et al., 2021).

The majority of the prior literature incorporates industry fixed effects into regression analyses aimed at identifying predictors of stock price crash risk, yet these effects are seldom elaborated upon. Both structural factors such as the competitive landscape, distribution channels and product differentiation, as well as operational factors such as litigation risk, resource dependencies, and sensitivity to shifts in government regulation, are all examples of factors which to a large extent depend on a firms' industry membership. There is a bulk of research supporting that industry membership can determine firm performance (e.g. Schmalensee, 1985), as well as research on how different industry characteristics impact stock performance (e.g. Hou and Robinson, 2006). It stands to reason that a similar association between industry membership and stock price crash risk might exist. While this area is largely unexplored, Li and Zang (2019) discover that competitive pressure augments crash risk using a sample from U.S. markets spanning 1998-2009, while Li and Luo (2020)

[1] Note that since we follow this definition specifically, the literature review specifically focuses on firm-specific crash risk, and disregards issues surrounding market-wide stock price crash literature and jump literature, as these are considered out of our purview.

reveal a negative correlation between higher industry competition and crash risk in the Chinese market These findings strengthen the proposition that industry-level effects have an impact on crash risk. Therefore, in this paper, we contribute to the existing literature by further examining whether a firm's industry influences stock price crash risk in U.S. stock markets.

3. Hypotheses

Previous research has revealed that news and events specific to a particular firm within an industry can influence the stock prices of other firms within the same industry through intra-industry effects. Lang & Stulz (1992) demonstrate that, on average, bankruptcy announcements result in decreased stock prices for industry competitors, and Slovin et al. (1991) report significant intra-industry effects stemming from bids to take firms private. Given that there are multiple factors that may impact stock price crash risk which are contingent upon a firm's industry membership, we put forth the following hypotheses:

> **H1: A firm's stock price crash risk is influenced by its industry, with variations in magnitude across industries.**

> **H2: The differences in magnitude can, at least in part, be attributed to industry characteristics.**

Potential industry-level factors that can impact crash risk can be broadly divided into two categories: (1) operational industry characteristics and (2) structural industry characteristics. Differences in crash risk across industries may be influenced by shared operational characteristics among firms within each industry. For instance, operational characteristics such as Research & Development intensity, labor intensity and resource dependency may contribute to certain industries being more susceptible to certain types of stock price crashes. On the other hand, structural characteristics refer to factors that determine crash risk through inter-firm interactions within an industry. While both categories have potential relevance for further research, this paper will focus specifically on structural characteristics. We have selected three structural industry characteristics that could potentially have implications for stock price crash risk: (1) Industry concentration, (2) Industry growth, and (3) Industry profitability.

Industry concentration is a characteristic that has been examined in asset pricing literature as a proxy for product market competition. For instance, industry concentration has been found to impact stock returns (Hou and Robinson, 2006) and market risk-return (Melicher et al., 1976). Industry competition can influence managerial incentives (Karuna, 2007), and according to Datta et al. (2013), industry competitiveness is positively associated with the extent of earnings management. Previous research thus suggests that high levels of competition could create incentives for managers to delay the disclosure of bad news. Therefore, within the framework of Jin and Meyers' (2006) agency theory approach to stock price crash risk, firms operating in industries with high levels of product market competition could be expected to have a higher stock price crash risk. Supporting this notion, Li and Zang (2019) have found a positive relationship between competitive pressure and crash risk for the period 1998-2009. Hence, we propose the following hypothesis concerning industry concentration:

H3: Firms operating in industries with a high market concentration face a lower stock price crash risk.

Another industry characteristic that holds potential relevance is industry growth rate. Jin and Meyers' (2006) theory of accumulation of bad news could be applicable to high-growth industries, given that firms within these industries tend to rely heavily on financing. DuCharme et al. (2001) and Teoh et al. (1998) argue that earnings management, driven by managers' incentives to achieve a high offering price, may explain the underperformance of IPOs. Similarly, managers of firms in high-growth industries may be motivated to delay the release of negative news in order to minimize their cost of capital. On the other hand, several researchers have found a positive relationship between firms' tendency to use external financing and their tendency to disclose earnings forecasts, as a high level of disclosure can reduce the cost of equity capital (e.g. Frankel et al, 1995; Botosan, 1997). Thus, previous research has not provided a clear indication of the effect that dependence on external financing has on stock price crash risk in the context of bad news disclosure. However, other aspects of industry growth rate may also be of relevance to stock price crash risk. For example, high-growth industries are characterized by rapid changes in market conditions and a high degree of uncertainty. As industry conditions evolve over the course of the industry life cycle, in conjunction with changes in the growth rate (Abernathy and Utterback, 1978; Jovanovic and MacDonald, 1994), the potential risk of stock price crashes may be affected. For example, the level of turbulence in an industry varies depending on its stage in the life cycle (Agarwal and Gort, 1996). Accordingly, we formulate our fourth hypothesis as follows:

H4: Firms operating in industries with high industry growth rates face a higher stock price crash risk.

Lastly, industry profitability may serve as a relevant determinant of stock price crash risk. Industry profitability is related to multiple different competitive forces both within an industry and stemming from interrelated industries (Porter, 2008). Therefore, while industry profitability is linked to both industry growth rate and industry concentration (Porter, 2008; Clarke et al., 1984), industry profitability as a variable may capture other competitive forces that impact industry participants. Given that competitive forces can affect managerial incentives (Karuna, 2007), it follows that such forces are likely to influence the extent of bad news hoarding within an industry. Industries with lower profitability face higher competitive forces (Porter, 2008), potentially incentivizing managers in these industries to delay the disclosure of negative news, thereby increasing the stock price crash risk. Consequently, we propose the following hypothesis:

H5: Firms operating in industries with low profitability face a higher stock price crash risk.

4. Variables and Data

4.1. Data Sources and Sample Selection

Following the previous literature (e.g. Callen and Fang, 2015; Chang et al., 2017; and Kim et al., 2011a) daily stock return data as well as annual and monthly accounting data is obtained from the Center for Research in Security Prices (CRSP) and Compustat merged database. U.S. stock markets are used because of large data availability and to allow for comparisons with prior research articles on the subject. Our sample period includes data from 1984-2022, which is the largest sample period available for the CRSP/Compustat merged database on Wharton Research Data Services. Notably, because of lags in variable calculations the final sample starts at 1985.

The collected security data includes each security's respective Global Industry Classification Standard (GICS) sector and industry codes, which is the leading privately provided solution for industry classification (Kaustia and Rantala, 2021). The main analysis uses the GICS 6-digit Industry Group definition, in favor of the 2-, 4-, or 8-digit codes. This is to ensure unrelated stocks are not grouped together, while still maintaining a large sample group for each industry classification.[2]

Following Kim et al. (2011a), we exclude observations with non-positive book values, non-positive total assets, fiscal year-end prices of less than one dollar, and with fewer than 26 weeks of stock return data. Additionally, following Chang et al. (2017) we further exclude observations with insufficient information for constructing the crash risk measures, and winsorize all variables except the crash dummy at the 1st and 99th percentile to reduce the effect of outliers. With these data requirements our final sample consists of 173640 firm-year observations for 18695 U.S. firms for the period 1985-2022.

4.2. Crash Risk Variables

Following the previous literature on crash risk (e.g. Chen et al., 2001; Jin and Myers, 2006; and Hutton et al., 2009), we construct several measures of firm-specific crash risk for each fiscal year. The three measures used are a negative skewness measure (NSKEW), a binary crash dummy (CRASH), and a "down-to-up volatility"-measure (DUVOL). Using several measures will give a more comprehensive view of crash risk and make our results more robust, and these three specifically are the most commonly used (e.g. Callen and Fang, 2015; Chang et al., 2017; Chen et al., 2001; Jin and Myers, 2006; and Hutton et al., 2009). Notably, while commonly used side by side, Chen et al. (2001) note a high correlation between NSKEW and DUVOL, indicating that they largely represent the same information.

All three of these measures are calculated based on firm-specific weekly returns. Following Hutton et al. (2009), we calculate an estimate of firm-specific weekly returns using residuals from an ordinary least squares regression of market returns. The bulk of the current literature control for both market returns and industry returns (e.g. Callen and Fang, 2015; Chang et al., 2017; and Hutton et al., 2009). However, since this analysis aims to capture potential industry specific effects we only control for market returns when calculating firm-specific weekly returns, as such:

$$r_{j,t} = \alpha_j + \beta_{1j} r_{MKT,t-1} + \beta_{2j} r_{MKT,t} + \beta_{3j} r_{MKT,t+1} + \varepsilon_{j,t} \tag{1}$$

[2] When using the 6-digit GICS definition, no industry has less than 100 observations.

where $r_{j,t}$ is the return for stock j in week t, $r_{MKT,t}$ is the return of the CRSP value-weighted market index in week t, and $\varepsilon_{j,t}$ is the residual return for stock j in week t. Subsequently, we estimate firm-specific weekly returns as the natural logarithm of 1 plus the firm-specific weekly market residual, $ln(1+\varepsilon_{j,t})$. Residuals are used to ensure the return estimate captures firm-specific returns and not broad market movements. The logarithmic transformation is used to create a more symmetric distribution of the otherwise skewed error term $\varepsilon_{j,t}$, which allows us to define a crash as residual returns corresponding to a threshold number of standard deviations below the mean (Hutton et al., 2009). This estimate of firm-specific weekly returns is then used to calculate the crash risk measures.

Our first measure of stock price crash risk is negative skewness (NSKEW). A higher NSKEW value signifies a more left-skewed stock return distribution, which indicates the stock is more "crash prone" (Callen and Fang, 2015). Following Chen et al. (2001) we calculate it by taking the negative of the third moment of firm-specific weekly returns for each sample year, divided by the standard deviation of firm-specific weekly returns raised to the third power, as follows:

$$NSKEW_{i,t} = -\frac{n(n-1)^{3/2}\Sigma W_{i,t}^3}{(n-1)(n-2)(\Sigma W_{i,t}^2)^{3/2}} \tag{2}$$

where n is the number of observations for firm j in fiscal year t. Due to its wide usage the NSKEW measure will be our primary measure for univariate analyses.

Our second measure of stock price crash risk is a binary crash risk measure (CRASH). Following the definition of CRASH used by Hutton et al. (2009), the dummy variable equals 1 if a firm experiences one or more weeks where the firm-specific weekly returns fall 3.09 standard deviations below the mean weekly firm-specific return for that fiscal year, and 0 otherwise. A fall of 3.09 standard deviations is thus representative of a week where the stock experienced a crash, as defined by a substantially negative weekly return. The number 3.09 has been chosen to generate a frequency of 0.1% given normally distributed returns.

The third measure of stock price crash risk is down-to-up volatility (DUVOL). Similarly to the NSKEW measure, a higher DUVOL value is indicative of a more crash prone stock. Contrary to NSKEW however, it is less likely to be overly influenced by a handful of extreme days since it does not involve third moments. Following Chen et al. (2001), DUVOL is calculated by taking the log of the ratio of the standard deviation on down weeks and the standard deviation of up weeks. Down weeks are defined as weeks with returns below the yearly mean, up weeks as weeks with returns above the yearly mean. The calculation is specified as follows:

$$DUVOL_{i,t} = \frac{log(n_u-1)\Sigma_{DOWN} R_{i,t}^2}{(n_d-1)\Sigma_{UP} R_{i,t}^2} \tag{3}$$

4.3. Industry Characteristics Variables

Three industry characteristics have been selected for this section: industry concentration, industry profitability, and industry growth rate. All industry characteristics for the main analysis are calculated on the 6-digit GICS Industry level unless otherwise specified. We exclude any observations with negative sales from our dataset.

To measure industry concentration, we use the Herfindahl-Hirschman Index, which is a commonly used index for market concentration (e.g. Hertzel and Officer, 2012; Sapienza, 2004). It is an index on a scale from 0 to 10,000, where a higher value is representative of a more concentrated, less competitive industry. It is calculated by squaring the market share of each firm in the industry and then summing the resulting numbers. This is done for each GICS industry using annual revenue data from the CRSP/Compustat merged database. In our analysis the variable is divided by 10,000, giving it a value between 0 and 1, to allow for easier interpretation of regression results.

For industry growth rate, we sum annual revenues for all firms within an industry, and calculate the percentage change in industry revenues for each year.

Several different measures of industry profitability exist, but we choose to use profit margin in our regression as it is largely affected by the competitive forces within an industry. Dou et al. (2021) argue that profit margins reflect competitive intensity, as industry players compete by undercutting the profit margins of rivals. We calculate a weighted average industry profitability by dividing the sum of total industry net income by the sum of total industry revenue. A weighted average measure is used in order to avoid the industry profitability measure being impacted by a few small companies with poor performance.

4.4. Control Variables

A number of variables will be controlled for in the regressions as prior research has demonstrated that these variables affect stock price crash risk. The selected control variables follow those used by Chang et al. (2017), who in turn follow the prior literature on crash risk.[3] Data from the CRSP/Compustat merged database has been used to construct these control variables. The crash risk measure NSKEW is included as a lagged control variable since stock return skewness has been shown to be time consistent (Chen et al., 2001). Stock return volatility, past stock returns, and past stock turnover are integrated in the regression as they have been shown to be positively associated with stock price crash risk (Chen et al., 2001). Following Chang et al., stock return volatility (SIGMA) is calculated as the standard deviation of firm-specific weekly returns for each fiscal year in the dataset, past stock returns (RET) is calculated as 100 times the mean of firm-specific weekly returns for each fiscal year in the dataset, and past stock turnover (DTURN) is calculated as average monthly stock turnover over the current fiscal year minus those over the previous fiscal year. Additionally, firm size and market-to-book ratio are controlled for as they have been documented as positively correlated with crash risk (Chen et al., 2001; Hutton et al., 2009). Firm size (SIZE) is calculated as the natural logarithmic of the market value of equity, and market-to-book ratio (MB) is the ratio of the market value of equity over the book value of equity. Furthermore, leverage and return on assets have been found negatively correlated with crash risk, and will thus be controlled for. Leverage (LEV) is defined as the ratio of long-term debt over the book value of total assets. Return on assets (ROA) is calculated as the ratio of income before extraordinary items divided by total assets. A full list of variable definitions can be found in Appendix A.

[3] One control variable used by Chang et al. (2017) as well as a lot of the other literature on stock price crash risk is discretionary accruals, as described by Hutton et al. (2009). However, this variable has been excluded from this analysis. This is because not all literature includes this variable (e.g. Chen et al., 2001), and there is a general divergence between what control variables are used by the top articles on this subject. It consequently seemed reasonable to exclude this specific variable due to its complicated calculation, as it seemed out of the scope of a BSc thesis to do such complex calculations for a control, non-key variable.

4.5. Descriptive Statistics

Table 1 reports the sample distribution by year for key variables. The key variables are the crash risk measures NSKEW, CRASH and DUVOL, as well as the industry characteristics Herfindahl-Hirschman Index (HHI), annual industry profitability, and annual industry growth rate. As expected, the crash risk measures are generally higher during crises, even though the measures are calculated using market residuals (see Section 4.2. and Equation 1). This can be seen for the 1987 Black Monday crash, the early 2000s dot-com bubble crash, and the 2008 to 2009 financial crisis. The values during these times are high relative to their values before and after the crisis for all the crash risk variables. Notably, the same pattern is not seen for the 2020 Covid-19 crash. This may be due to the market-wide nature of the crash, and the quick market recovery. Nonetheless, these results suggest that the crash risk measures do indeed capture moments where stocks experience crashes.

Turning to the industry characteristics, HHI seems to have been somewhat stable during the sample period. Mean Industry Profitability is generally higher during the 2000s, peaking in 2021 and only showing negative results in 2001, assumingly as a result of the 2000s dot-com bubble crash. Industry growth rate seems to have lower values post-2000, with less values above 10% and several negative values. The negative values coincide with both the financial crisis and the Covid-19 crisis, indicating industries may contract in terms of revenue during crises.

Table 2 presents summary statistics and a Pearson correlation matrix of both key and control variables. The NSKEW, CRASH and DUVOL variables have means that are slightly higher than the bulk of prior literature, likely due to our crash risk measures being calculated without controlling for industry effects (see Section 3.2. for details).[4] Additionally, the past stock returns control variable (RET) seems to be a bit higher than in several other papers on stock price crash risk (e.g. Chang et al., 2017; Hutton et al., 2009). This could possibly be due to this paper including a longer time period. In the correlation matrix, the NSKEW and DUVOL variables show a very high correlation (0.952), indicating they largely represent the same information, as was noted by Chen et al. (2001). None of the crash risk measures or industry characteristic variables show an extremely high correlation with any of the control variables, with the strongest correlation being industry profitability and the stock return volatility measure SIGMA with a negative 0.165 correlation. Out of the three industry characteristics, industry profitability is the one with the highest correlation with the control variables, with three variables having a correlation with an absolute value above 0.1. Additionally, none of the industry characteristic variables show an extremely high correlation with any of the other industry characteristic variables. This is beneficial as too high correlations within the regression increases the risk of multicollinearity leading to misleading results. However, the fact that the industry concentration measure and industry profitability measure show very low correlation should be kept in mind throughout the analysis, as they are two different measures of competitive forces and one could therefore expect higher competition. No clear pattern between the crash risk variables and the industry characteristic variables is obvious in the correlation matrix.

[4] Parts of this difference could also be attributable to other factors, such as different datasets due to different timeframes.

Table 1: Distribution by Year For Key Variables

Table 1 reports the mean of the variables NSKEW, CRASH, DUVOL, the Herfindahl-Hirschman Index (HHI), Annual Profitability, and Annual Growth Rate, for each separate year in the sample data. The sample is stock data from the Center for Research in Security Prices (CRSP) and Compustat merged database, for the period 1985-2022. NSKEW is a crash measure that measures negative skewness, CRASH is a binary crash risk dummy measuring crash weeks, and DUVOL measures down-to-up volatility. Full crash risk definitions can be found in Appendix A.1. The HHI measure is a market concentration index, calculated by summing the squares of the market shares of all firms in an industry. Industry profitability is calculated as the average of the profit margin of all firms within an industry in the sample. Industry Growth Rate is defined as the annual percentage change in industry revenues, where industry revenue is the sum of annual revenues for all firms within an industry in the sample. All tabulated industry characteristics are based on the GICS 6-digit Industry Group classification.

Year	N	NSKEW Mean	CRASH Mean	DUVOL Mean	HHI Mean	Profitability Mean	Growth Mean
1985	3637	0.211	0.245	-0.020	0.102	0.033	0.062
1986	3677	0.234	0.239	-0.014	0.099	0.032	0.055
1987	4268	0.302	0.273	0.006	0.099	0.043	0.144
1988	4020	-0.110	0.177	-0.144	0.103	0.045	0.102
1989	3919	0.073	0.225	-0.064	0.104	0.039	0.058
1990	3707	-0.224	0.299	-0.131	0.103	0.036	0.090
1991	3916	0.096	0.198	-0.046	0.101	0.028	0.048
1992	4578	0.244	0.271	-0.005	0.097	0.028	0.077
1993	5187	0.203	0.235	-0.014	0.086	0.038	0.117
1994	5601	0.178	0.220	-0.013	0.083	0.053	0.133
1995	5915	0.163	0.234	-0.034	0.080	0.054	0.163
1996	6113	0.175	0.210	-0.029	0.076	0.054	0.143
1997	6368	0.268	0.234	-0.005	0.078	0.052	0.130
1998	6429	0.323	0.288	0.028	0.080	0.054	0.103
1999	6009	0.028	0.213	-0.096	0.083	0.058	0.084
2000	5525	0.175	0.254	-0.020	0.086	0.046	0.154
2001	5344	0.0004	0.209	-0.088	0.088	-0.003	0.010
2002	4943	0.390	0.312	0.105	0.091	0.010	0.010
2003	5111	0.210	0.233	0.014	0.093	0.058	0.081
2004	5034	0.281	0.304	0.033	0.094	0.074	0.124
2005	4853	0.281	0.294	0.041	0.095	0.086	0.090
2006	4817	0.230	0.280	0.020	0.098	0.090	0.120
2007	4570	-0.006	0.414	-0.043	0.102	0.078	0.049
2008	4103	0.483	0.378	0.138	0.099	0.018	-0.010
2009	4415	0.279	0.278	0.009	0.102	0.055	-0.070
2010	4171	0.005	0.206	-0.060	0.102	0.084	0.094
2011	4016	0.083	0.220	-0.022	0.102	0.082	0.075
2012	4001	0.092	0.206	-0.020	0.101	0.078	0.041
2013	4060	0.086	0.247	-0.028	0.099	0.089	0.038
2014	4080	-0.026	0.200	-0.085	0.094	0.093	0.052
2015	4244	0.218	0.315	0.039	0.096	0.084	-0.031
2016	4027	0.043	0.276	-0.040	0.093	0.082	-0.009
2017	3983	0.133	0.313	0.002	0.093	0.086	0.092
2018	3851	-0.095	0.328	-0.076	0.094	0.098	0.057
2019	3920	0.082	0.266	-0.026	0.096	0.092	0.040
2020	4188	0.015	0.224	-0.082	0.097	0.049	-0.022
2021	4185	-0.020	0.198	-0.074	0.092	0.117	0.195
2022	2858	0.216	0.253	0.051	0.101	0.101	-0.030
All	**173640**	**0.148**	**0.256**	**-0.020**	**0.093**	**0.059**	**0.075**

Table 2: Summary Statistics and Correlation Matrix

Table 2 reports summary statistics (Panel A) and a Pearson correlation matrix (Panel B) for the sample. The sample is stock data from the Center for Research in Security Prices (CRSP) and Compustat merged database, for the period 1985-2022. The selected crash risk measures are NSKEW, CRASH, and DUVOL. The NSKEW crash measure is defined as the negative of the third moment of firm-specific weekly returns for each sample year divided by the standard deviation of firm-specific weekly returns raised to the third power. CRASH is a binary crash risk measure which equals 1 if a stock experiences one or more weeks where firm-specific weekly returns fall 3.09 standard deviations below the mean weekly firm-specific return for that fiscal year, and 0 otherwise. DUVOL is defined as the log of the ratio of the standard deviation on "down days" and the standard deviation of "up days". Down days are defined as days with returns below the yearly mean, up days as days with returns above the yearly mean. Three industry characteristics have been identified, namely market concentration as measured by the Herfindahl-Hirschman Index (HHI), Annual Profitability, and Annual Growth Rate. Other variables are control variables. Full list of variable definitions can be found in Appendix A. In Panel B, the stars (*) represent statistically significant results at the 1% (***), 5% (**) and 10% (*) level.

Panel A: Summary Statistics

Statistic	N	Mean	St. Dev.	Min	Q1	Q3	Max
NSKEW	173,640	0.148	1.152	-2.425	-0.430	0.451	5.135
CRASH	173,640	0.256	0.437	0	0	1	1
DUVOL	173,640	-0.020	0.451	-1.019	-0.301	0.200	1.582
RET	173,640	0.629	1.637	-1.941	-0.119	0.882	10.767
SIGMA	173,640	0.098	0.105	0.019	0.046	0.108	0.779
SIZE	173,640	5.702	2.253	1.151	4.011	7.263	11.240
MB	173,640	3.050	4.252	0.296	1.116	3.152	31.001
LEV	173,640	0.169	0.178	0.000	0.009	0.282	0.701
ROA	173,640	-0.016	0.187	-1.012	-0.011	0.064	0.273
DTURN	173,640	0.003	0.088	-0.339	-0.019	0.020	0.421
HHI	173,640	0.093	0.074	0.023	0.045	0.116	0.433
Industry Growth Rate	173,640	0.075	0.136	-0.323	0.004	0.140	0.553
Industry Profitability	173,640	0.059	0.071	-0.225	0.029	0.095	0.249

Panel B: Correlation Matrix

	NSKEW	CRASH	DUVOL	RET	SIGMA	SIZE	MB	LEV	ROA	DTURN	HHI	Ind. Growth	Ind. Prof.
NSKEW	1.000												
CRASH	0.637***	1.000											
DUVOL	0.952***	0.628***	1.000										
RET	-0.148***	-0.094***	-0.206***	1.000									
SIGMA	0.004	0.042***	-0.055***	0.777***	1.000								
SIZE	0.122***	0.055***	0.164***	-0.105***	-0.246***	1.000							
MB	0.021***	0.012***	0.005*	0.122***	0.058***	0.163***	1.000						
LEV	-0.009***	-0.014***	0.011***	-0.048***	-0.053***	0.160***	0.069***	1.000					
ROA	0.045***	0.011***	0.077***	-0.121***	-0.305***	0.268***	-0.169***	0.065	1.000				
DTURN	0.034***	0.055***	0.017***	0.195***	0.161***	0.023***	0.062***	0.001	-0.009***	1.000			
HHI	-0.005*	0.008***	-0.006**	0.024***	0.030***	0.002	0.062***	0.046***	-0.023***	0.005*	1.000		
Ind. Growth	0.010***	-0.009***	-0.000	-0.007**	-0.018***	-0.056***	0.052***	-0.042***	-0.010***	0.013***	0.026***	1.000	
Ind. Prof.	0.002	-0.000	0.022***	-0.061***	-0.165***	0.159***	-0.003	0.017***	0.142***	0.003	0.030***	0.117***	1.000

*p<0.1; >**p<0.05; >***p<0

5. Empirical Analysis

5.1. Measuring Industry Fixed Effects

The first part of our empirical analysis attempts to determine whether there are significant industry effects on stock price crash risk, and if so, in which industries this effect is the strongest. To do this, we will first perform a univariate analysis through plotting crash risk over time, in order to visualize differences between sectors' crash risk. Secondly, we will regress our crash risk measures on sector and industry dummies, and establish which industries have the strongest and weakest relationships with crash risk. Then, crash risk is plotted over time for the industries with the strongest and weakest relationships, in order to visually examine if any clear patterns are present over time.

5.1.1. Univariate Analysis: Yearly Crash Risk For Different Sectors

We begin by plotting crash risk for different sectors, for the sample period 1985-2022. To do this, crash risk is proxied by an annual sector NSKEW, calculated for each sector by summing each firm's NSKEW multiplied by the firm's market share. Sector classifications come from 11 different 2-digit GICS sectors, which have been used in favor of 6-digit industries to allow for easier interpretation. Additionally, the lines have been smoothened to facilitate interpretation.

Figure 1 reports the result of sectors' crash risk plotted over time. In the figure, the sectors' NSKEW values do not necessarily move together, which could indicate there are some industry fixed effects. It does also seem as if sectors show more extreme NSKEW values at similar times, whether positive or negative. This can be seen in the divergence of the sectors' crash risk data around the mid-1990s. However, whether or not these patterns are because of sector effects specifically cannot be determined by this graph. These effects could also be attributable to sectors being correlated with factors that affect crash risk, such as firm size, return on assets, and leverage. Thus, to more definitively determine what effect sectors and industries have on crash risk, multivariate regressions that control for such variables must be done.

5.1.2. Multivariate Analysis: Measuring Industry Fixed Effects

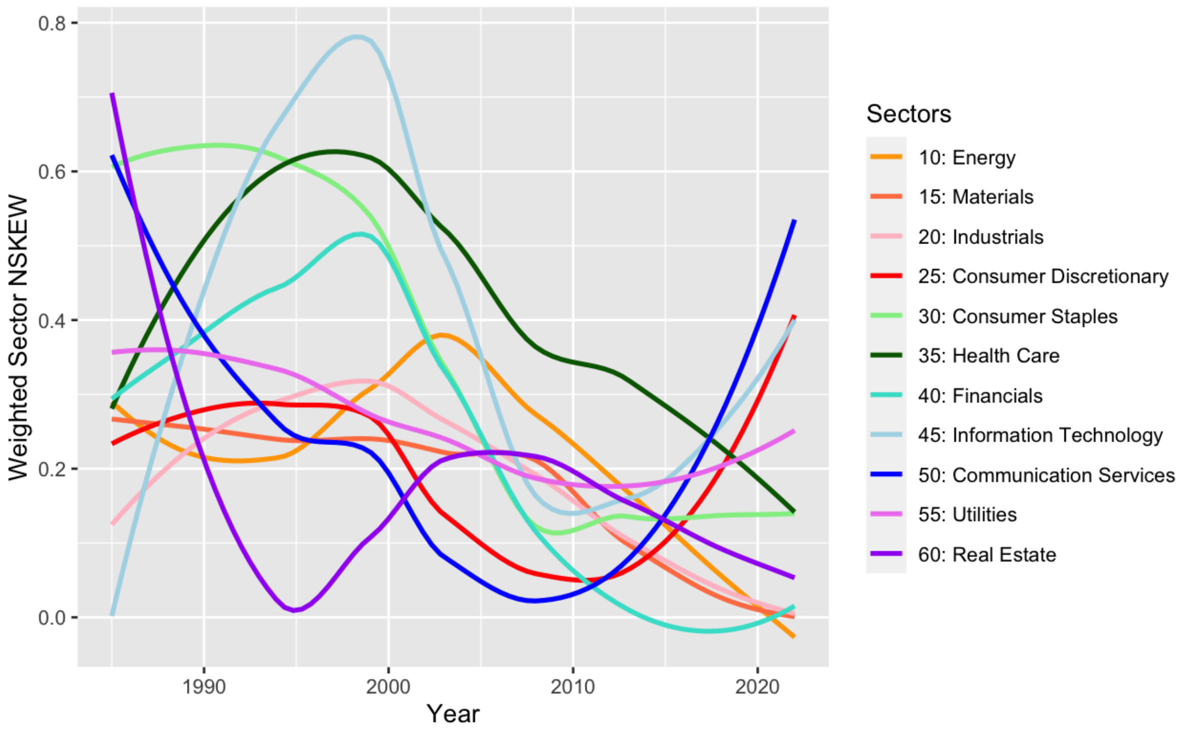
To further establish the existence of possible industry effects on stock price crash risk, we will try to quantify this effect while controlling for firm-specific characteristics that have an effect on stock price crash risk. To do this, three regressions on the crash risk measures NSKEW, CRASH and DUVOL are done. The regressions are defined as follows:

$$\begin{gathered} CRASHMEASURE_{i,t} = \beta_0 + \beta_1 NSKEW_{i,t-1} + \beta_2 SIGMA \\ +\beta_3 RET_{i,t-1} + \beta_4 DTURN_{i,t-1} + \beta_5 SIZE_{i,t-1} + \beta_6 MB_{i,t-1} \\ +\beta_7 LEV_{i,t-1} + \beta_8 ROA_{i,t-1} + YR_t + IND_{i,t} + \varepsilon_{i,t} \end{gathered} \tag{4}$$

In the regressions, $CRASHMEASURE_{i,t}$ refers to our crash measures NSKEW, CRASH and DUVOL, *i* denotes firm, *t* denotes the year, YR_t denotes year fixed effects, and

Figure 1: Crash Risk For Sectors Over Time

Figure 1 presents the average industry crash risk over the period 1985-2022 for different sectors. The sample is stock data from the Center for Research in Security Prices (CRSP)/Compustat merged database for the period 1985-2022. Sector crash risk is proxied by Industry NSKEW, which is calculated as an average of annual crash risk for all firms within a sector, weighted against the market share of each firm in the sector. NSKEW's firm-specific definition can be found in Appendix A.1. The sector dummies are based on the Global Industry Classification Standard (GICS) 2-digit sector classification.



$IND_{i,t}$ denotes industry fixed effects.[5] The industry fixed effects are based on the 2-digit sector or 6-digit industry GICS classification. All independent variables except the year fixed effects and industry fixed effects are lagged by one year. Following Chang et al. (2017), a logit model is used for the CRASH regression due to the dependent variable being binary, and an ordinary least squares (OLS) regression when NSKEW and DUVOL are the dependent variables.

First, regressions that use 2-digit GICS sector classifications as industry dummies are performed. The aim of these regressions is to determine whether or not the sectors show significant effect on crash risk, and at what magnitudes. 2-digit sectors have been used in favor of 6-digit industries for this regression, as the large number of 6-digit industries in the GICS system makes a regression with industry classifications tough to tabulate effectively.

Table 3 shows the output of the regressions with 2-digit sector classifications as proxies for industry fixed effects. Importantly, when making a regression using industry dummies, one of the dummies is omitted, in this case *Sector 10: Energy*. This is because its inclusion becomes redundant. Instead, the results for the other sectors will be interpreted relative to *Sector 10: Energy*.

---

[5] The reason the industry dummies include a time subscript component is that some of the GICS industries change over time. For example, the industry *Real Estate Management & Development* had the code 404030 until it was discontinued in 2016. Then it had the code 601020, and then got the new name 602010.

Table 3: Regression Results With Sector Fixed Effects

Table 3 reports the regression results for the regressions on stock price crash risk, which are regressed on sector dummies and selected control variables. The sample is stock data from the Center for Research in Security Prices (CRSP)/Compustat merged database for the period 1985-2022. The selected crash risk measures are NSKEW, CRASH, and DUVOL. The NSKEW crash measure is defined as the negative of the third moment of firm-specific weekly returns for each sample year divided by the standard deviation of firm-specific weekly returns raised to the third power. CRASH is a binary crash risk measure which equals 1 if a stock experiences one or more weeks where firm-specific weekly returns fall 3.09 standard deviations below the mean weekly firm-specific return for that fiscal year, and 0 otherwise. DUVOL is defined as the log of the ratio of the standard deviation on "down days" and the standard deviation of "up days". Down days are defined as days with returns below the yearly mean, up days as days with returns above the yearly mean. Control variable definitions can be found in Appendix A.2. The sector dummies are based on the Global Industry Classification Standard (GICS) 2-digit sector classification.

	Dependent variable:		
	NSKEW	CRASH	DUVOL
	OLS	*Normal*	*OLS*
	(1)	(2)	(3)
Lagged NSKEW	0.036*** (0.002)	0.012*** (0.001)	0.017*** (0.001)
Lagged SIGMA	-1.341*** (0.050)	-0.321*** (0.019)	-0.620*** (0.019)
Lagged RET	0.093*** (0.003)	0.023*** (0.001)	0.040*** (0.001)
Lagged DTURN	0.058* (0.032)	0.019 (0.012)	0.018 (0.012)
Lagged SIZE	0.074*** (0.001)	0.013*** (0.001)	0.033*** (0.001)
Lagged MB	0.007*** (0.001)	0.003*** (0.0003)	0.003*** (0.0003)
Lagged LEV	-0.167*** (0.016)	-0.043*** (0.006)	-0.051*** (0.006)
Lagged ROA	0.094*** (0.017)	0.044*** (0.006)	0.085*** (0.007)
Sector 15: Materials	0.005 (0.016)	0.026*** (0.006)	0.001 (0.006)
Sector 20: Industrials	0.054*** (0.014)	0.065*** (0.005)	0.019*** (0.005)
Sector 25: Consumer Discretionary	0.061*** (0.014)	0.075*** (0.005)	0.022*** (0.005)
Sector 30: Consumer Staples	0.072*** (0.017)	0.067*** (0.007)	0.023*** (0.007)
Sector 35: Health Care	0.108*** (0.014)	0.085*** (0.005)	0.030*** (0.005)
Sector 40: Financials	0.078*** (0.013)	0.057*** (0.005)	0.035*** (0.005)
Sector 45: Information Technology	0.099*** (0.014)	0.089*** (0.005)	0.032*** (0.005)
Sector 50: Communication Services	0.037* (0.020)	0.043*** (0.007)	0.002 (0.008)
Sector 55: Utilities	0.084*** (0.020)	0.018** (0.008)	0.057*** (0.008)
Sector 60: Real Estate	0.039* (0.020)	0.017** (0.008)	0.044*** (0.008)
Observations	173,639	173,639	173,639
R^2	0.048		0.063
Adjusted R^2	0.048		0.063
Note:			*p<0.1; >**p<0.05; >***p<0.01
			Year fixed effects is in the regression but has been omitted from the output.

Almost all sectors show a significant positive effect on all crash risk measures at the 5 or 1 percent significance level. However, *Sector 15: Materials* only show significant results on the CRASH dummy, and Sector 50: Communication Services only show significant results on the NSKEW measure and CRASH dummy. Since all results are positive relative to the baseline dummy of *Sector 10: Energy*, this would suggest that the energy sector has the weakest relationship with stock price crash risk among the sectors. We can also see that the magnitude of the effect relative to the energy sector differs. For example, in the regression with the NSKEW crash measure, *Sector 50: Communication Services* has an effect of 0.037, and *Sector 35: Health Care* has an effect of 0.108, which is almost three times as big of an effect relative to the energy sector. Similar differences are found between the strongest and weakest effect for the CRASH and DUVOL measures.

In the NSKEW regression, the sectors with the effects of the highest magnitude are *Sector 35: Health Care*, *Sector 45: Information Technology*, and *Sector 55: Utilities*; in the CRASH regression *Sector 45: Information Technology*, *Sector 35: Health Care*, and *Sector 25: Consumer Discretionary*; and in the DUVOL regression *Sector 55: Utilities*, *Sector 60: Real Estate*, and *Sector 40: Financials*. On the flipside, in the NSKEW regression, the sectors with the effects of the lowest magnitude are *Sector 10: Energy*,[6] *Sector 50: Communication Services*, and *Sector 60: Real Estate*; in the CRASH regression *Sector 10: Energy*, *Sector 60: Real Estate*, and *Sector 55: Utilities*; and in the DUVOL regression *Sector 10: Energy*, *Sector 20: Industrials*, and *Sector 25: Consumer Discretionary*.

It thus seems that between the three crash risk measures some sectors repeatedly show effects of high magnitude on stock price crash risk, such as the health care sector and the information technology sector. Others repeatedly show effects of low magnitude on stock price crash risk, such as the energy sector. Some sectors show ambiguous results between the different crash risk measures, such as the real estate sector, the utilities sector, and the consumer discretionary sector. This ambiguity is due to the fact that the different crash risk measures capture crash risk differently. Overall, the regressions suggest that sector dummies relate significantly to stock price crash risk at differing magnitudes.

Untabulated regressions using the 6-digit industry classifications instead of the 2-digit sector regression provide additional insights. For this regression, the reference industry dummy that is omitted is *Industry 101010: Energy Equipment & Services.* Contrary to the 2-digit regression, with 6-digit industry dummies we see that not all industries have significant effects, and the magnitude differs more greatly. For example, for the NSKEW measure only 25 out of 83 industries show significant results at the 5 percent level. Additionally, more industries show negative effects on stock price crash risk compared to the reference dummy. The divergence between the 2-digit and 6-digit analysis is likely because using the more granular 6-digit industry classification ensures there is a smaller chance of stocks with very different characteristics being grouped together.

To quantify and tabulate how this effect differs across 6-digit industries, the coefficient estimates and p-values for each industry binary from the 6-digit regressions are extracted and ordered. This is done to determine which industry dummies are most strongly and weakly related to stock price crash risk. Table 4 demonstrates the results from the regression with NSKEW as the crash risk measure. Panel A reports the industries that have the strongest positive relationship to crash risk (i.e. higher crash risk industries), and Panel B reports the industries with the strongest negative relationship or weakest positive relationship to crash risk (i.e. lower crash risk industries). Only industries that have a significant effect (defined as a p-value lower than 0.05) on the crash risk measure are included.

In Panel A of Table 4 we can see that out of the top 10 industries with effects of the highest magnitude at a significant level, six out of ten industries' GICS codes start with either 35, 40 or 55. This indicates they are classified as part of *Sector 35: Health Care*, *Sector 40: Financials*, or *Sector 55: Utilities*. In Panel B we can see that out of the top 10 industries with effects of the lowest magnitude at a significant level, nine out of ten industries belong to either *Sector 10: Energy*, *Sector 20: Industrials*, *Sector 25: Consumer Discretionary*, or *Sector 40: Financials.* Based on these panels, it seems some sectors include industries with significant effects of *either* high and low magnitudes. For example, *Sector 35: Health Care* only has industries in Panel A, indicating it primarily includes industries that relate strongly

[6] Note that while this sector has been omitted from the output, we can still discern that this sector has the effect of the lowest magnitude for all crash risk measures. This is because all other sectors show positive effects, indicating all others are positive *relative* to the energy sector.

Table 4: Industries With The Strongest Industry Effects

Table 4 reports the industries with the strongest and weakest effect on stock price crash risk, as determined by a regression where crash risk is regressed on industry dummies and selected control variables (see Appendix A.3. for control variables). The sample is stock data from the Center for Research in Security Prices (CRSP)/Compustat merged database for the period 1985-2022. Crash risk is proxied by negative skewness (NSKEW), which is defined as the negative of the third moment of firm-specific weekly returns for each sample year divided by the standard deviation of firm-specific weekly returns raised to the third power. The industry dummies are based on the Global Industry Classification Standard (GICS) 6-digit industry classification. The P-value for industry *101010: Energy Equipment & Services* is not available since it is the reference point for the other dummies.

Panel A: Top 10 Industries With Effects of Greatest Magnitude on NSKEW Measure, Given $p<0.05$.

GICS Code	Industry Name	Coefficient	P-Value
402040	Mortgage Real Estate Investment Trusts (REITs)	0.150	0.004
351030	Health Care Technology	0.132	0.009
351020	Health Care Providers & Services	0.120	0.000
253020	Diversified Consumer Services	0.119	0.005
551040	Water Utilities	0.118	0.026
203040	Ground Transportation (New Name)	0.106	0.005
451020	IT Services	0.102	0.002
551020	Gas Utilities	0.101	0.012
402020	Consumer Finance	0.100	0.024
301010	Consumer Staples Distribution & Retail (New Name)	0.095	0.008

Panel B: Top 10 Industries With Effects of Lowest Magnitude on NSKEW Measure, Given $p<0.05$.

GICS Code	Industry Name	Coefficient	P-Value
404030	Real Estate Management & Development (Discontinued 2016)	-0.191	0.013
201050	Industrial Conglomerates	-0.166	0.008
251020	Automobiles	-0.140	0.012
151020	Construction Materials	-0.106	0.046
254010	Media (Discontinued 2018)	-0.080	0.014
201010	Aerospace & Defense	-0.072	0.026
101020	Oil, Gas & Consumable Fuels	-0.063	0.021
101010	Energy Equipment & Services	0	N/A*
401010	Banks	0.054	0.035
252030	Textiles, Apparel & Luxury Goods	0.069	0.030

*The p-value for industry *101010: Energy Equipment & Services* is not available due to the industry being used as the reference in the regression.

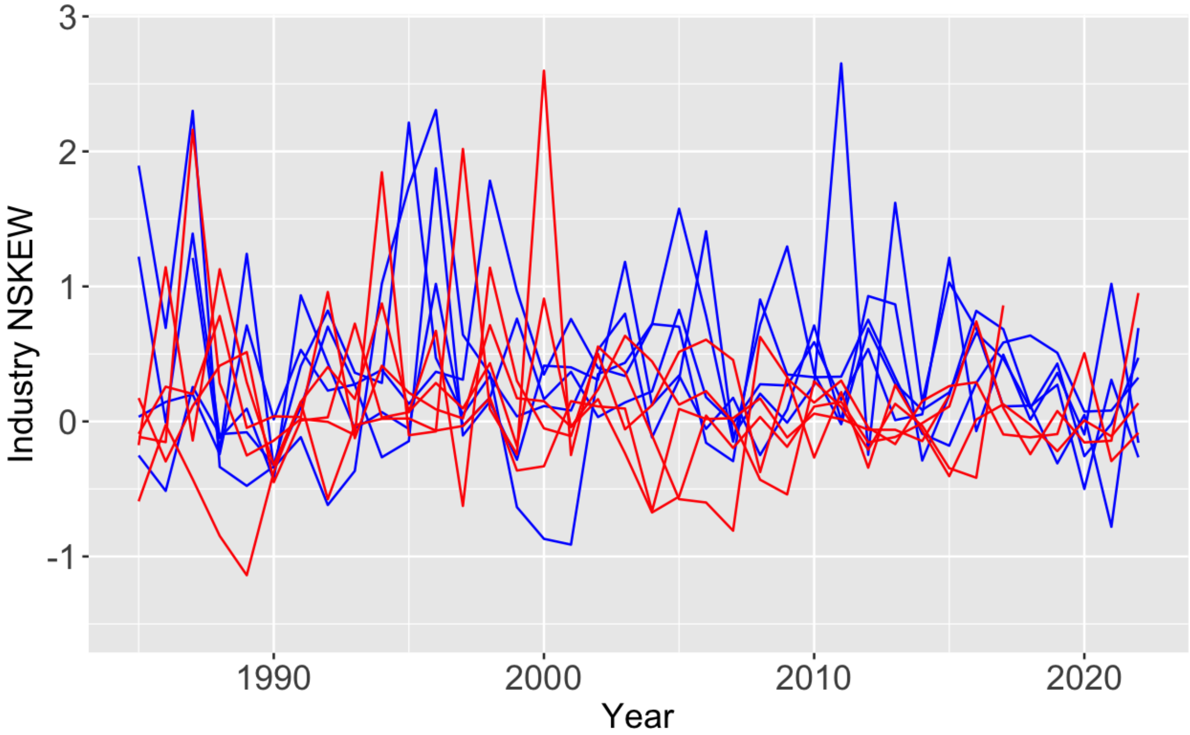
to crash risk. Additionally, *Sector 10: Energy* only has two industries in total, both of which are present in Panel B, indicating they relate weakly to crash risk. However, some sectors include industries that have significant effects of both high and low magnitudes, such as *Sector 40: Financials* and *Sector 25: Consumer Discretionary*, who have industries in both Panel A and Panel B. Table A in Appendix B reports the same results but for the CRASH and DUVOL crash risk measures. The overall conclusion remains the same no matter the crash risk measure.

For further insights into how industry relates to crash risk, we look at how this effect has developed over time. Figure 2 presents the weighted average industry NSKEW for selected industries over the sample period of 1985-2022. Similarly to in Figure 1, yearly average NSKEW for each industry is calculated as a weighted average measure of annual

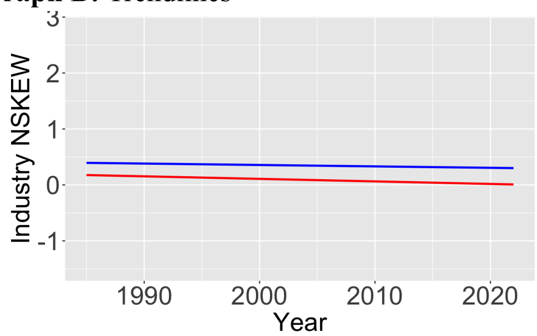
Figure 2: Crash Risk For Selected Industries Over Time

Figure 2 presents the average industry crash risk over the period 1985-2022 for selected industries. Industry crash risk is proxied by Industry NSKEW, which is calculated as an average of annual crash risk for all firms within an industry, weighted against the market share of each firm in the industry. The industries are 6-digit GICS industries, and those chosen are those that have displayed the strongest and weakest effect on crash risk in multivariate analyses. The blue lines represent industries that have had a strong effect on price risk: 402040: Mortgage Real Estate Investment Trusts (REITs), 351030: Health Care Technology, 351020: Health Care Providers & Services, 253020: Diversified Consumer Services, and 401020: Thrifts & Mortgage Finance (Discontinued). The red lines represent industries that have had a weak effect on price risk:404030: Real Estate Management & Development (Discontinued 2016), 201050: Industrial Conglomerates, 251020: Automobiles, 254010: Media (Discontinued 2018), and 201010: Aerospace & Defense. Panel A plots the variables. Panel B and C plots timelines for the strong and weak effect industries, for the whole period and for 1984-2000 and 2001-2022 separately.

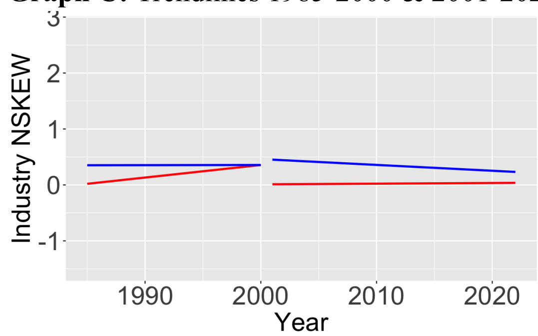
Graph A: All 10 Industries Plotted



Graph B: Trendlines



Graph C: Trendlines 1985-2000 & 2001-2022



crash risk for each industry, with the market share of each firm within each industry used as weight. The selected industries are the top five industries with effects of the highest and lowest magnitudes, as reported in Table 4. Specifically, the blue lines are the five industries with effects of the highest magnitude as reported in Panel A of Table 4: *402040: Mortgage Real Estate Investment Trusts (REITs)*, *351030: Health Care Technology*, *351020: Health Care Providers & Services*, *253020: Diversified Consumer Services*, and *551040: Water*

Utilities. The red lines are the five industries the effects of the lowest magnitudes as reported in Panel B of Table 4: *404030: Real Estate Management & Development (Discontinued 2016)*, *201050: Industrial Conglomerates*, *251020: Automobiles*, *151020: ConstructionMaterials,* and *254010: Media (Discontinued 2018)*. In Graph A, average industry NSKEW is plotted for each of these ten industries. In Graph B a trendline for the industries with effects of high magnitude (blue) and for the industries effects of low magnitudes (red) are plotted, for the entire sample period 1985-2022. In Graph C similar trendlines are plotted but divided into the periods 1985-2000 and 2001-2022.

In Graph A in Figure 2, no trend is immediately apparent. Crash risk for the different industries seem to be inconsistent, with high peaks for both industries that are strongly and weakly related to crash risk. However, roughly around the year 2000 there seems to be some divergence between the industries with strong effects (blue) and the industries with weak effects (red). Specifically, it seems the blue lines are, on average, slightly above the red lines. Based on this, it seems the industry effects may have changed over the years, with a shift roughly around the year 2000.

In Graph B, the trend lines for the industries with strong effects (blue) and the industries with weak effects (red) are plotted. The blue trend line is slightly above the red line, with similar slopes. This seems to indicate that the industries with strong effects (blue) do indeed have slightly higher crash risk.

In Graph C, trend lines are once again plotted, but this time divided into two periods, namely 1985-2000 and 2001-2022. These periods were chosen based on the observation that there seemed to be some shift roughly around the year 2000 in Graph A. For the period 1985-2000, the trendlines start separate and then converge around the year 2000. For the period 2001-2022 a similar pattern is detected. This pattern insinuates that the industry effects for the sample industries was getting smaller up until the year 2000, and then the effects diverged. Additionally, the red lines are always below the blue lines, insinuating that those industries have always had industry effects of lower magnitude.

The results from Table 3, Table 4, and Figure 2 lead us to three conclusions. First, the sector/industry a stock belongs to significantly relates to the stock's crash risk.[7] Note that this conclusion does not necessarily mean that simply belonging to a specific industry will lead to a higher or lower crash risk, as this effect could be due to some omitted industry characteristic. Additionally, this effect could also be due to firm-specific characteristics that are correlated with industries. To ensure this is not the case for any of the firm-specific characteristics we have included as control variables, we have conducted Variance Inflation Factor (VIF) tests for all the above mentioned regressions. No variable in any of the regressions have VIF scores above 5, which indicates multicollinearity is not a grave issue. The second conclusion is that the significance and magnitude of this effect differs between different sectors/industries. The third conclusion is that this effect seems to differ between the periods 1985-2000 and 2001-2022. This brings us to the second part of our analysis, in which we will regress crash risk on certain industry characteristics in order to determine some of the dynamics that may be underlying the results in this section.

Notably, we also check whether the control variables are consistent with the prior literature for the regressions in this section. The results are largely consistent, such as positive results for lagged NSKEW, past stock returns (RET), stock turnover (DTURN), market-to-book ratio (MB), firm size (SIZE), and negative results for leverage (LEV). There are some slight divergences, such as us having negative results for stock volatility (SIGMA), positive results for return on assets (ROA), and some insignificant results for stock turnover

[7] Note that this conclusion rests on the assumption that the results do not entirely come from issues with the regression, such as multicollinearity.

(DTURN). This could be due to calculation differences,[8] or the fact that we include a dataset with more recent data than the bulk of prior literature. Additionally, the fact that DTURN shows insignificant results in some of the regressions may be due to the fact that it was calculated using common stocks instead of total stocks due to limited data availability.

5.2. Industry Characteristics Analysis

The second part of our analysis explores possible explanations for the industry effects that were established in the previous section. Specifically, we examine whether three structural industry characteristics relate to stock price crash risk, namely industry concentration, industry growth rate, and industry profitability. This is done in two parts. First, we conduct a univariate analysis where we look at how the industry characteristics relate to crash risk. Second, we conduct a multivariate analysis through a plurality of regressions where crash risk is regressed on the industry characteristics.

5.2.1. Univariate Analysis: Mean NSKEW For Different Industry Characteristics Deciles

We begin our analysis on the relationship between the industry characteristics and stock price crash risk by plotting our crash risk measure NSKEW against industry concentration, industry growth rate, and industry profitability respectively. The dataset is divided into deciles based on these three industry characteristics lagged by one year, and mean NSKEW is calculated for each decile. Figure 3 reports the plots of the mean NSKEW values against the deciles for each industry characteristic. Graph A of Figure 3 shows market concentration against NSKEW. No specific trend is apparent, which suggests industry concentration may not be a significant determinant of stock price crash risk. Graph B shows industry growth rate against NSKEW. The plot suggests an increasing trend in stock price crash risk as industry growth increases. Graph C shows industry profitability against NSKEW. This plot suggests a positive relationship between crash risk and industry profitability between the first and eighth decile, however this relationship appears to shift between the eighth and tenth decile, suggesting that crash risk decreases as industry profitability becomes very high. These findings suggest that industry growth and industry profitability may be determinants of stock price crash risk. Market concentration, on the other hand, does not appear to be related to stock price crash risk.

Based on the observation in Section 5.1.2 where we found that the industry effects seem to differ between the periods 1985-2000 and 2001-2022, we also conduct the same univariate analysis as above but with the data divided into those two periods. The results are presented in Figure 4. In Panel A, HHI deciles are plotted against NSKEW. Similarly to in Figure 3, no pattern is immediately apparent, except possibly that very low market concentration was seemingly related to high crash risk in 1985-2000. In Panel B, industry growth rate deciles are plotted against NSKEW. Once again the patterns largely resemble those in Figure 3, where we see a slight upward trend in both time periods. This could indicate a higher industry growth rate leads to higher crash risk, no matter the time period. In Panel C, industry profitability deciles are plotted against NSKEW. When comparing the two plots, the patterns are very different. In the period 1985-2000 it seems there is a distinct upwards trend, yet in 2001-2022 the plot is quite stable across the deciles. This could suggest that a higher profitability led to higher crash risk in the period 1985-2000, but that this effect

[8] For example, the fact that we only control for market returns and not industry returns in our calculation of our crash risk measures (see Equation 1) could affect these results.

was not present in the period 2001-2022. However, to positively determine whether this is indeed the case, multivariate analyses that control for other factors should be conducted.

5.2.2 Multivariate Analysis: Determining The Effect of Industry Characteristics

To determine whether our three structural industry characteristics have any effect on stock price crash risk, we regress the industry-level characteristics on our stock price crash risk measures while controlling for firm-level characteristics. The regressions are specified as follows:

$$\begin{gathered} CRASHMEASURE_{i,t} = \beta_0 + \beta_1 HHI_{i,t} + \beta_2 GROWTH_{i,t-1} + \beta_3 PM_{i,t-1} \\ + \beta_4 NSKEW_{i,t-1} + \beta_5 SIGMA_{i,t-1} + \beta_6 RET_{i,t-1} + \beta_7 DTURN_{i,t-1} + \beta_8 SIZE_{i,t-1} \\ + \beta_9 MB_{i,t-1} + \beta_{10} LEV_{i,t-1} + \beta_{11} ROA_{i,t-1} + IND_{i,t} + YR_t + \varepsilon_{i,t} \end{gathered} \tag{5}$$

Similarly to Equation 4, i denotes firm, t denotes the year, and YR_t denotes year fixed effects. However, as compared to Equations 4 these regressions include the addition of $HHI_{i,t-1}$, $GROWTH_{i,t-1}$ and $PM_{i,t-1}$. These capture the industry-level Herfindahl-Hirschman Index, industry growth rate and industry profitability, respectively, for each firm. These industry characteristic measures have been calculated for each industry based on 6-digit GICS industry codes.

The results of these regressions can be seen in Table 5. In Panel A the crash risk variables are regressed on all three industry characteristic variables simultaneously. In Panel B we only include the industry concentration HHI index, in Panel C we only include the industry growth rate variable, and in Panel D we only include the industry profitability variable. For brevity the control variables are untabulated in the latter three panels.

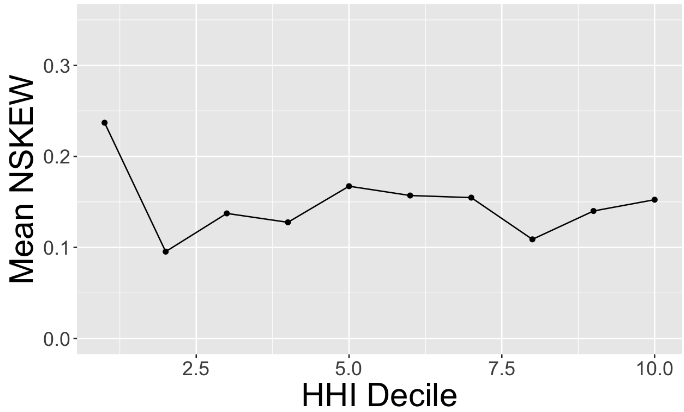
As seen in the table, the industry concentration variable HHI shows no significant results, suggesting that industry concentration does not affect the crash risk for stocks. This contradicts the prior literature, as significant negative results have been found on industry concentration in Chinese markets (Li and Luo, 2020) and significant positive results have been found in U.S. markets for the period 1998-2009 (Li and Zhan, 2019). It is, however, worth noting that the method used in this paper for constructing the Herfindahl-Hirschman index variable has limitations. We base our HHI-estimates on Compustat data, which only covers publicly listed firms. This means our HHI-estimates do not take private firms into account, which may affect our results (Ali et al., 2008). Hence, the results for our market concentration variable should be interpreted with caution, and we can not draw any definitive conclusion regarding the absence of a relationship between industry concentration and stock price crash risk. For this reason, we examine the impact of substituting Compustat-based concentration measures for the U.S. Census Bureau-based measures in Section 5.3.

The industry growth rate variable shows positive significant results at the 1 percent level for all regressions. These results suggest that stocks that belong to industries with relatively high industry growth rates are more likely to experience a stock crash. Further research is needed to determine the mechanisms underlying this result. As hypothesized, it could be an effect of high levels of industry turbulence (Agarwal & Gort, 1996). It could also be an effect of information asymmetry mechanisms due to a reliance on external financing. There are thus two different possibilities for how industry growth rate could affect bad news hoarding. Hence, future research could examine how dependence on external financing impacts stock price crash risk, and whether managers are incentivized to delay bad news

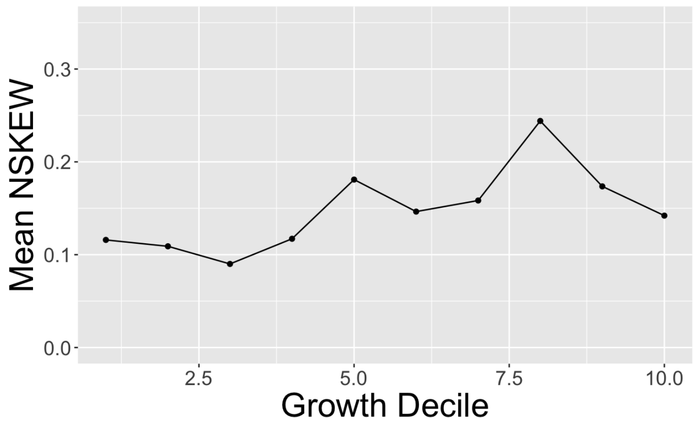
Figure 3: Plot of Crash Risk Against Industry Characteristics Deciles

Figure 3 shows the mean values of crash risk for industry characteristics deciles. Crash risk is proxied by NSKEW, which is defined as the negative of the third moment of firm-specific weekly returns for each sample year divided by the standard deviation of firm-specific weekly returns raised to the third power. The sample is stock data from the Center for Research in Security Prices (CRSP)/Compustat merged database for the period 1985-2022. Graph A shows mean NSKEW plotted against deciles of lagged industry concentration, as measured by the Herfindahl-Hirschman Index (HHI). Graph B shows the mean of NSKEW plotted against deciles of lagged industry growth. Graph C shows mean NSKEW plotted against deciles of lagged industry profitability. The HHI measure is a market concentration index, calculated by summing the squares of the market shares of all firms in an industry. Industry profitability is calculated as the average of the profit margin of all firms within an industry in the sample. Industry Growth Rate is defined as the annual percentage change in industry revenues, where industry revenue is the sum of annual revenues for all firms within an industry in the sample. All tabulated industry characteristics are based on the GICS 6-digit Industry Group classification.

Graph A: Industry Concentration Deciles



Graph B: Growth Deciles



Graph C: Profitability Deciles

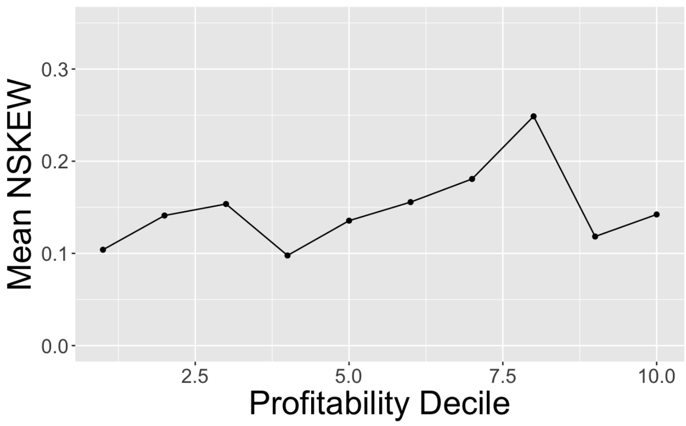
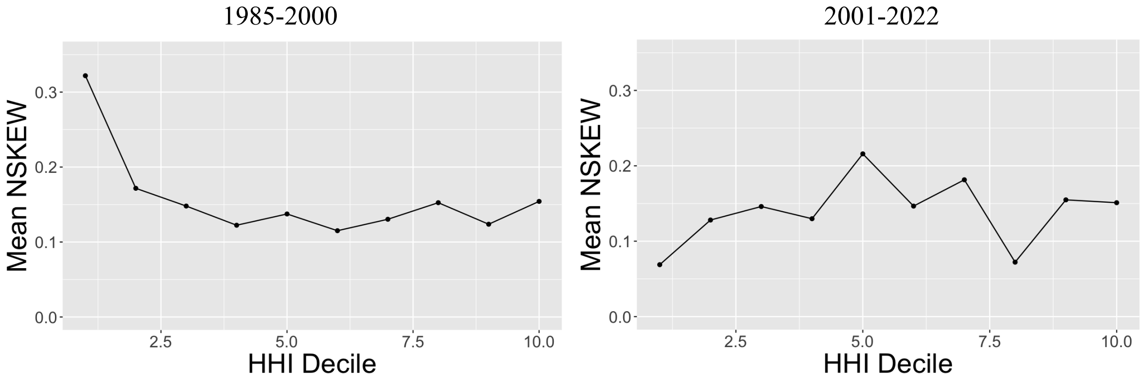


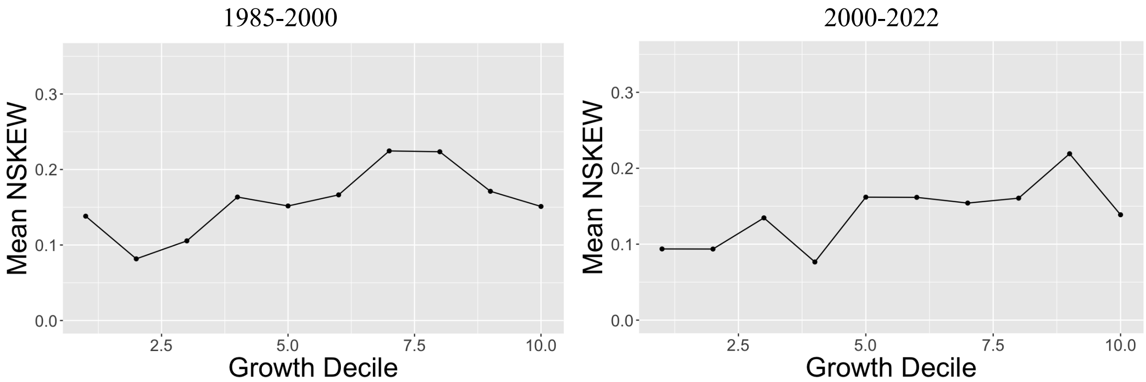
Figure 4: Crash Risk Against Industry Characteristics Deciles, Two Periods

Figure 4 shows the mean values of crash risk for industry characteristics deciles. Crash risk is proxied by NSKEW, which is defined as the negative of the third moment of firm-specific weekly returns for each sample year divided by the standard deviation of firm-specific weekly returns raised to the third power. The data has been divided into two periods, 1985-2000 and 2001-2022. The sample is stock data from the Center for Research in Security Prices (CRSP)/Compustat merged database for the period 1985-2022. Graph A shows mean NSKEW plotted against deciles of lagged industry concentration, as measured by the Herfindahl-Hirschman Index (HHI). Graph B shows the mean of NSKEW plotted against deciles of lagged industry growth. Graph C shows mean NSKEW plotted against deciles of lagged industry profitability. The HHI measure is a market concentration index, calculated by summing the squares of the market shares of all firms in an industry. Industry profitability is calculated as the average of the profit margin of all firms within an industry in the sample. Industry Growth Rate is defined as the annual percentage change in industry revenues, where industry revenue is the sum of annual revenues for all firms within an industry in the sample. All tabulated industry characteristics are based on the GICS 6-digit Industry Group classification.

Graph A: Industry Concentration Deciles



Graph B: Industry Growth Rate Deciles



Graph C: Industry Profitability Deciles

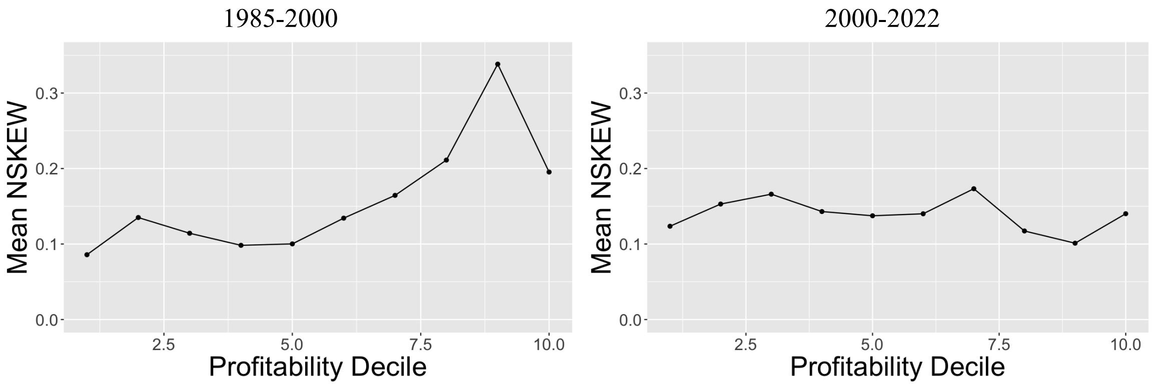


Table 5: Regression Results With Industry Characteristics

Table 5 reports the results of the regressions where three crash risk measures (NSKEW, CRASH, and DUVOL) are regressed on three different industry characteristics (the Herfindahl-Hirschman Index (HHI), Annual Profitability, and Annual Growth Rate). In Panel A all three industry characteristics are included in the same regression. In Panel B only HHI is included, in Panel C only Annual Industry Profitability, and in Panel D only Annual Industry Growth Rate. For brevity the control variables are untabulated for Panel B, C and D. The sample is stock data from the Center for Research in Security Prices (CRSP) and Compustat merged database, for the period 1985-2022. NSKEW is a crash measure that measures negative skewness, CRASH is a binary crash risk dummy measuring crash weeks, and DUVOL measures down-to-up volatility. Full crash risk definitions can be found in Appendix A.1. The HHI measure is a market concentration index, calculated by summing the squares of the market shares of all firms in an industry. Industry profitability is calculated as the average of the profit margin of all firms within an industry in the sample. Industry Growth Rate is defined as the annual percentage change in industry revenues, where industry revenue is the sum of annual revenues for all firms within an industry in the sample. All tabulated industry characteristics are based on the GICS 6-digit Industry Group classification.

	Dependent variable:		
	NSKEW	CRASH	DUVOL
	OLS	*Normal*	*OLS*
Panel A: All Industry Characteristics			
Lagged HHI	-0.053 (0.055)	-0.009 (0.021)	-0.031 (0.021)
Lagged Ind. Growth Rate	0.131*** (0.022)	0.038*** (0.009)	0.059*** (0.009)
Lagged Ind. Profitability	-0.072 (0.050)	-0.011 (0.019)	-0.050** (0.019)
Lagged NSKEW	0.035*** (0.002)	0.012*** (0.001)	0.016*** (0.001)
Lagged SIGMA	-1.340*** (0.051)	-0.323*** (0.019)	-0.618*** (0.020)
Lagged RET	0.093*** (0.003)	0.023*** (0.001)	0.040*** (0.001)
Lagged DTURN	0.055* (0.032)	0.019 (0.012)	0.016 (0.012)
Lagged SIZE	0.075*** (0.002)	0.013*** (0.001)	0.034*** (0.001)
Lagged MB	0.007*** (0.001)	0.002*** (0.0003)	0.003*** (0.0003)
Lagged LEV	-0.156*** (0.017)	-0.041*** (0.007)	-0.054*** (0.007)
Lagged ROA	0.082*** (0.017)	0.037*** (0.007)	0.076*** (0.007)
Panel B: Only HHI*			
Lagged HHI	-0.048 (0.055)	-0.008 (0.021)	-0.029 (0.021)
Panel C: Only Industry Growth Rate*			
Lagged Ind. Growth Rate	0.125*** (0.022)	0.037*** (0.008)	0.056*** (0.009)
Panel D: Only Industry Profitability*			
Lagged Ind. Profitability	-0.029 (0.049)	0.002 (0.019)	-0.031 (0.019)

*For brevity, control variables are included in the regression but omitted from the output.of Panel B, C, and D. Additionally, year and industry fixed effects are included in all the regressions.

releases (Teoh et al., 1998; DuCharme et al., 2001), or to increase disclosure (Frankel et al, 1995; Botosan, 1997), in order to minimize cost of capital.

Industry profitability does not show significant results other than negative results in the DUVOL regression when the other industry-level characteristics are also included. This indicates that while relatively high industry profitability could possibly lead to stock being less prone to crashes, the results are not robust and no definitive conclusions can be drawn. This also suggests that the seemingly positive trend that was seen in the univariate analysis in

Figure 3 may be based on industry profitability having correlations with other factors that have positive effects on crash risk, as opposed to industry profitability itself having positive effects on crash risk.

However, when conducting Variance Inflation Factor (VIF) tests on the regressions, we find that the industry dummies have a VIF-score above 5 in the regressions where all industry characteristics are included simultaneously (i.e. Panel A of Table 5). This indicates multicollinearity may affect the results. However, when the industry characteristics are included one-by-one in the regressions (i.e. Panel B, C, and D), no VIF score is above 5. This indicates multicollinearity is not a grave issue in those regressions. Since the results in Panel A do not differ largely from Panel B, C and D, the above analysis of the regressions still holds.

Based on the observation in Section 5.1.2 where we found that the industry effects seem to differ between the periods 1985-2000 and 2001-2022, we also conduct the same regressions as above but divide the dataset into the two time periods. Additionally, since multicollinearity may be an issue if all variables are included in the regression simultaneously, we only include one industry characteristic per regression. The results of these regressions are in Table 6.

Similarly to Table 5, Table 6 shows no significant results for the industry concentration HHI index. However, both the results for industry growth rate and industry profitability differ in Table 6 as compared to Table 5. While industry growth rate showed positive significant results for the entire sample period in Table 5, it only has significant results for the latter time period (2001-2022) for the NSKEW and DUVOL variables in Table 6. The CRASH variable still shows positive significant results for both time periods. These results suggest that the effect we found in the entire sample period (Table 5) is not present in the period 1985-2000 when isolated.

Industry profitability did not have significant results for the entire sample period in Table 5, but in Table 6 we see that the variable has significant negative effects on all crash risk measures in the period 1985-2000. These results suggest that during the period 1985-2000 a relatively high industry profitability was related to lower stock price crash risk. Additionally, significant positive results are found for the period 2001-2022 for the NSKEW measure, which could suggest that higher industry profitability leads to higher crash risk for the period 2001-2022. However, as these results were only observed for one of the three crash risk measures, no robust conclusions can be drawn about the period 2001-2022. This confirms the conclusion of the univariate analysis in Figure 3 in Section 5.2.1., where we found that profitability seemed to relate to profitability in 1985-2000 but not as clearly in 2001-2022. However, the negative coefficients of the regressions for the period 1985-2000 contradicts the assumption of the effect being positive due to the seemingly positive trend of the plot in Figure 3. Due to this contradiction, the seemingly positive trend in the univariate analysis is likely driven by industry profitability being correlated with other factors we control for. As noted in Section 4.5., industry profitability has correlations above an absolute value of 0.1 with three of the control variables, namely stock volatility (SIGMA), firm size (SIZE), and return on assets (ROA). However, when conducting VIF tests on the regressions, no variable in the regressions with profitability has VIF-score above 5, which indicates multicollinearity is not a grave issue. The only regressions which have a variable with a VIF-score above 5 are the 2001-2022 HHI regressions, but since no conclusions are drawn from those regressions this is not an issue for our interpretation of the results.

Collectively our results suggest that relatively high industry growth rates overall lead to higher crash risk for stocks, but that this effect was not present in the time period 1985-2000. Additionally, we find that higher industry profitability led to lower stock price crash risk in the period 1985-2000, but find no conclusive results for the period 2001-2022. No connection has been found between industry concentration and stock price crash risk.

Table 6: Regression Results With Industry Characteristics, Two Periods

Table 6 reports the results of regressions where the crash risk measures are regressed on three different industry characteristics (the Herfindahl-Hirschman Index (HHI), Annual Profitability, and Annual Growth Rate). The first regression includes data from the period 1985-2000, and the second regression includes data from the period 2001-2022. The sample is stock data from the Center for Research in Security Prices (CRSP) and Compustat merged database. The primary dependent variables are the crash risk measures NSKEW and CRASH. Full definition of these variables can be found in Appendix A.1. The HHI measure is a market concentration index, calculated by summing the squares of the market shares of all firms in an industry. Industry profitability is calculated as the average of the profit margin of all firms within an industry in the sample. Industry Growth Rate is defined as the annual percentage change in industry revenues, where industry revenue is the sum of annual revenues for all firms within an industry in the sample. All tabulated industry characteristics are based on the GICS 6-digit Industry Group classification.

Panel A: Regressions with NSKEW as crash risk measure*

	Dependent variable:	
	NSKEW	
	1985-2000	2001-2022
Regression 1: Lagged HHI	0.133 (0.093)	0.033 (0.084)
Regression 2: Lagged Industry Growth Rate	0.025 (0.037)	0.104*** (0.028)
Regression 3: Lagged Industry Profitability	-0.259** (0.102)	0.134** (0.056)

Panel B: Regressions with CRASH as crash risk measure*

	Dependent variable:	
	CRASH	
	1985-2000	2001-2022
Regression 1: Lagged HHI	-0.004 (0.033)	0.005 (0.035)
Regression 2: Lagged Industry Growth Rate	0.037*** (0.013)	0.043*** (0.012)
Regression 3: Lagged Industry Profitability	-0.101*** (0.036)	0.022 (0.023)

Panel C: Regressions with DUVOL as crash risk measure*

	Dependent variable:	
	DUVOL	
	1985-2000	2001-2022
Regression 1: Lagged HHI	0.045 (0.035)	0.028 (0.033)
Regression 2: Lagged Industry Growth Rate	0.018 (0.014)	0.053*** (0.011)
Regression 3: Lagged Industry Profitability	-0.121*** (0.039)	-0.022 (0.022)

*For each panel, three separate regressions are tabulated per time period, one for each industry characteristic. Thus, there are a total of six regressions per panel, and 18 separate regressions in the entire table. For brevity, the control variables are untabulated.

A topic for further research could be to determine the reasons behind this difference between the two time periods. It is difficult to state any preliminary hypothesis, as many factors may be behind the discrepancy. For example, Regulation Fair Disclosure was adopted in the US at the end of year 2000, and the dot-com bubble began to burst in 2001. Hence, additional research is needed in order to establish potential causes.

5.3. Robustness Tests

To ensure the robustness of our results, several different tests with different variable definitions have been done. The results of these tests are reported in Table 7, with only the coefficients tabulated for conciseness.[9]

First, we consider the robustness of the stock price crash risk measures. Since we already use three distinct measures of crash risk (CRASH, NSKEW, and DUVOL) our results are seemingly already quite robust to this measure. However, even though the CRASH dummy is widely used in prior literature (e.g. Chang et al., 2017; Hutton et al., 2009), it can be considered less informative than other measures of crash risk due to its binary nature (Chang et al., 2017). Consequently, we check its robustness by using a CRASH measure which counts the *total* number of "crash weeks" during a year, as opposed to a binary measure equaling 1 if a stock experiences one or more "crash weeks" during a year. Since the variable is no longer binary an ordinary least squares regression is used instead of a logit regression. The results of this analysis can be found in Panel A of Table 7.

Second, we examine whether the choice of using 6-digit GICS industry classifications affect the effect the chosen industry characteristics (HHI, Industry Growth Rate, and Industry Profitability) have on crash risk. The industry characteristics measures have thus been recalculated using 2-digit, 4-digit, and 8-digit GICS classifications, which are then regressed on our crash risk. For brevity, only the results for the CRASH dummy and NSKEW measures are tabulated.[10] The results from these regressions are tabulated in Panel B, C, and D of Table 7.

Third, we consider whether using GICS industry classifications specifically alter the results. We thus recalculate the industry characteristics using the Standard Industry Classification (SIC) standard, as well as the North American Industry Classification Standard (NAICS) to define our industries. The results with these measures are tabulated in Panel E and F of Table 7. Untabulated results show that these industry classifications also have significant industry fixed effects, as discussed in Section 5.1.

Fourth, we examine whether alternative definitions of our industry characteristics affect the results. We use industry employment growth, defined by the annual percentage change in industry employment level, as an alternative measure of industry growth rate. Employment growth is a variable that has been widely used in research on firm and industry growth (e.g. Kumar, 1985; Evans, 1987; Birley et al., 1990). Janssen (2009) argues that employment growth and sales growth are the most commonly used growth measures in research, but that they are determined by distinct factors and thus capture different aspects of growth. We obtain employment growth data based on North American Industry Classification Standard (NAICS) industries from the U.S. Bureau of Labor Statistics for privately owned firms. This data is only available for the years 2018-2021 and thus results in a limited sample size of 14858 observations. The results of our regression using industry employment growth as our measure for industry growth is tabulated in Panel G of Table 7.

---

[9] In addition to the tests we conducted, we also wanted to control for firm fixed effects as it seems customary in the prior literature on crash risk (e.g. Chang et al., 2017). However, due to constraints in computational power with the large dataset we were unable to do this.

[10] These two are chosen as NSKEW is arguably the most commonly used stock price crash risk measure, and it has a higher correlation with the DUVOL measure as compared to the CRASH dummy, making the inclusion of CRASH more likely to provide results that diverge from the NSKEW measure

Table 7: Robustness Tests

Table 7 reports the results of a plurality of robustness tests. The sample is stock data from the Center for Research in Security Prices (CRSP) and Compustat merged database, for the period 1985-2022. The primary dependent variables are the crash risk measures NSKEW and CRASH. Full definition of these variables can be found in Appendix A.1. The alternative CRASH measure is defined similarly to CRASH, but is not binary and instead counts the total number of crash weeks. Definitions of the dependent variables HHI, Industry Growth Rate, and Industry Profitability can be found in Appendix A.2. Panel A reports the results for when the industry characteristics are regressed on the alternative crash risk measure. For Panel B, C, and D the industry characteristics have been calculated using 2-, 4-, and 8-digit GICS codes respectively. For Panel E Standard Industry Classifications are used. For Panel F the North American Industry Classification System has been used. For Panel G alternative profitability measures are used as independent variables in the place of Industry Profitability. For Panel G an alternative industry growth measure is used in the place of Industry Growth Rate.

	Dependent variable:		
	Alternative CRASH	NSKEW	CRASH
	Normal	*OLS*	*Normal*
Panel A: Alternative Crash Risk Measure			
Lagged HHI	-0.012 (0.022)		
Lagged Ind. Growth Rate	0.039*** (0.009)		
Lagged Ind. Profitability	-0.022 (0.020)		
Panel B: 2-Digit GICS Classification			
Lagged HHI		-0.583** (0.276)	0.030 (0.106)
Lagged Ind. Growth Rate		0.306*** (0.038)	0.063*** (0.015)
Lagged Ind. Profitability		-0.206** (0.099)	0.023 (0.038)
Panel C: 4-Digit GICS Classification			
Lagged HHI		-0.208** (0.105)	0.052 (0.040)
Lagged Ind. Growth Rate		0.225*** (0.031)	0.052*** (0.012)
Lagged Ind. Profitability		-0.324*** (0.070)	-0.048* (0.027)
Panel D: 8-Digit GICS Classification			
Lagged HHI		-0.019 (0.027)	0.000 (0.010)
Lagged Ind. Growth Rate		0.089*** (0.017)	0.023*** (0.007)
Lagged Ind. Profitability		-0.014 (0.037)	0.017 (0.014)
Panel E: Standard Industry Classification			
Lagged HHI		-0.018 (0.016)	0.004 (0.006)
Lagged Ind. Growth Rate		0.061*** (0.015)	0.017*** (0.006)
Lagged Ind. Profitability		-0.082** (0.038)	-0.006 (0.015)
Panel F: North American Industry Classification System			
Lagged HHI		-0.005 (0.012)	-0.001 (0.004)
Lagged Ind. Growth Rate		0.078*** (0.012)	0.021*** (0.005)
Lagged Ind. Profitability		-0.071** (0.031)	-0.027** (0.012)
Panel G: Alternative Industry Growth Rate Measure			
Average Industry Employee Growth		0.046 (0.138)	N/A*
Panel H: Alternative Industry Concentration Measure			
Compustat Based CR4		-0.098*** (0.033)	-0.015 (0.012)
Panel I: Alternative Profitability Measures			
Median Industry Return on Assets		0.324*** (0.090)	0.088*** (0.034)
Median Industry Return on Equity		0.241*** (0.049)	0.049*** (0.019)
Median Industry Return on Equity Employed		0.395*** (0.064)	0.092*** (0.024)

*The values for CRASH risk have been omitted because the sample size was limited, leading to the control variables not showing significant results. The control variables were however still significant for the NSKEW variable, which is why it is still included.

Compustat-based measures of industry concentration only consider public firms, and consequently may be poor proxies for actual industry concentration (Ali et al., 2008). Consequently, we check the robustness of our result through substituting it with a dataset from the U.S. Census Bureau, with the Herfindahl-Hirschman index (HHI) measure and the alternative industry concentration measure four-firm concentration ratios (CR4) for SIC-industries between 1985-1992, and regress our crash risk variables on these alternative measures. Due to the relatively short timespan, this data results in limited sample sizes of only 2774 and 2747 respectively. In the regressions we found that the control variables showed no significant results, and have thus chosen to disregard those regressions as the results are likely unreliable. We also calculate a Compustat-based measure of four-firm concentration ratio (CR4) based on GICS 6-digit industries and perform a regression with this measure as our industry concentration variable. The results from this regression is shown in Panel H of Table 7.

We also test whether using different measures of industry profitability gives the same results as our main regression analyzes. To do this, we use three alternative profitability definitions: return on assets (ROA), return on equity (ROE) and return on capital employed (ROCE). We obtain monthly firm-level ratios from Compustat and calculate annual industry mean values for all three measures. The results of using each measure, respectively, for industry profitability are shown in Panel I of Table 7.

Generally the results of the robustness tests align with the results reported in our main analysis. The main divergence is found for the alternative industry concentration measure, where CR4 calculated based on Compustat data is used instead of HHI. Here we find negative significant results for the NSKEW measure at the 1 percent level, but no results for the CRASH measure. Untabulated results of the CR4 variable in a regression on DUVOL also finds negative significant results (a coefficient of -0.061, significant at the 1 percent level with a standard error of 0.013). These results contradict our main analysis, where we found no significant results. Instead these results suggest that a higher industry concentration could lead to lower crash risk, as hypothesized in Section 3. This suggests there could be some effect from industry concentration, and that the proxy we use does not capture these effects. Thus, no definitive conclusions can be drawn from the results on industry concentration, but these results indicate this topic could be of interest for further research.

In line with our main analysis, industry growth rate consistently has positive and significant results for all different industry classifications, and for the alternative CRASH measure. However, the alternative industry growth rate measure (employee growth rate) shows no significant results. However, only 4 years of data was available for employee growth data. So while this measure does not show significant results, it does not mean the results from the main analysis are void, but rather that their robustness is not guaranteed by this measure. Further, Janssen (2009) argues that employment growth and revenue growth capture different aspects of growth rates for firms. This indicates that it could be possible that industry revenue growth significantly impacts stock price crash risk while industry employment growth does not.

The industry profitability variable shows ambiguous results, with some industry classifications getting negative significant results and some not getting significant results, just like in our main analysis. However, contrary to our main analysis the alternative profitability measures do all show significant positive results. This could indicate industry profitability has a significant positive effect on stock price crash risk, and that the variable in our main analysis does not capture this effect. However, when conducting VIF tests on these regressions the alternative industry profitability measures report VIF scores above 5, indicating multicollinearity could significantly impact these results. Thus, the results for industry profitability remain ambiguous.

6. Conclusion

This study investigates industry-level effects on stock price crash risk. Specifically, it examines whether there are industry fixed effects on crash risk, and if so, in which industries this effect is the strongest. Subsequently, three different industry characteristics are investigated as possible explanations for industry-level differences in crash risk. The chosen industry characteristics are industry concentration, industry growth rate, and industry profitability.

We find evidence of significant industry fixed effects at differing magnitudes, indicating different industries have differing likelihoods of crash risk. For example, industries in the health care sector are consistently related to relatively high crash risk, industries in the energy sector are consistently related to relatively low crash risk, and industries in the financial sector are related to both relatively high and low crash risk.

Having established that industry-level effects on crash risk exist at differing magnitudes, we examined the three industry characteristics' relationship to stock price crash risk. We find evidence that relatively high industry growth rates relate to higher crash risk. However, this effect was not found for the period 1985-2000 when isolated. Additionally, we find evidence that relatively high industry profitability was related to lower crash risk in the period 1985-2000. No such relationship can be conclusively determined for 2000-2022. These results are robust to different industry classifications and crash measures, and are significant even while controlling for factors known to affect stock price, such as past returns and stock volatility (Chen et al., 2001).

No significant results were found for our industry concentration proxy. However, as we got significant negative results when robustness testing the measure, we can not conclusively determine that no relationship exists between industry concentration and stock price crash risk.

This study complements the existing literature on stock price crash risk, as industry-level effects on stock price crash risk have been largely disregarded by the bulk of the prior literature on the subject. Thus, these results enrich our understanding of the influence of industry-level effects on future stock price crash risk, which has a material impact on investors' welfare.

While this study identified some possible factors that contribute to this effect, the area is worth researching further. For one, industry concentration could be further investigated using proxies other than the Herfindahl-Hirschman Index computed using Compustat data. Second, the divergence of industry effects between the two time periods 1985-2000 and 2001-2022 could be further examined. Such research could both investigate whether this effect is present in some of the factors known to affect stock price, and try to find the reasons for this divergence. Lastly, industry characteristics that have been disregarded in this study could be investigated. For example, other structural industry characteristics such as product differentiation could be investigated, or more operational industry characteristics such as resource dependency, capital intensity, or research and development intensity.

Appendix A: Variable Definitions

The parenthesized, capitalized parts of this section refer to the specific code attributed to a variable on the Center for Research in Security Prices (CRSP) and Compustat merged database, as downloaded from Wharton Research Data Services (WRDS).

A.1. Crash Risk Measures

CRASH: A dummy variable which equals 1 if a stock experiences one or more weeks where firm-specific weekly returns fall 3.09 standard deviations below the mean weekly firm-specific return for that fiscal year, and 0 otherwise. Firm specific weekly returns is defined as the natural logarithm of 1 plus the firm-specific weekly market residual, with the residual being from a regression that controls for market returns (see Equation 1).

DUVOL: The log of the ratio of the standard deviation on down weeks and the standard deviation of up weeks. Down weeks are defined as weeks with firm specific weekly returns below the yearly mean, up weeks as weeks with firm specific weekly returns above the yearly mean. Firm specific weekly returns is defined as the natural logarithm of 1 plus the firm-specific weekly market residual, with the residual being from a regression that controls for market returns (see Equation 1).

NSKEW: The negative of the third moment of firm-specific weekly returns for each sample year divided by the standard deviation of firm-specific weekly returns raised to the third power. Firm specific weekly returns is defined as the natural logarithm of 1 plus the firm-specific weekly market residual, with the residual being from a regression that controls for market returns (see Equation 1).

A.2. Industry Characteristic Variables

Growth Rate: The annual percentage change in industry revenues, where industry revenues is the sum of annual revenues (REVT) for all firms within an industry in the sample.

HHI: The Herfindahl-Hirschman Index, a market concentration index calculated by summing the squares of the market shares, calculated using net sales (SALES), in percent of all firms in an industry in the sample. Divided by 10000 to transform it from 0-10,000 to 0-1, to facilitate the interpretation of regressions.

Profitability: The weighted average profit margin of all firms in an industry within the sample. Calculated by dividing the sum of net income (NI) within the industry by the sum of revenue (REVT).

A.3. Control Variables

DTURN: Average monthly stock turnovers over the current fiscal year minus those over the previous fiscal year. Monthly stock turnover is calculated as the monthly trading volume (CSHTRMS) divided by the amount of shares outstanding (CSHOC). Since the number of shares outstanding is reported on a quarterly basis, the months in between are proxied with last quarter's reporting.

LEV: The ratio of long-term debt (DLTT) over the book value of total assets (AT).

MB: The ratio of the market value of equity over the book value of equity (CEQ). Market value of equity is calculated as a stock's annual closing price (PRCC_F) times the amount of common shares (CSHPRI)[11].

[11]Common shares were used in favor of total shares (common and preferred stocks) due to limited data availability.

RET: 100 times the mean of firm-specific weekly returns for each fiscal year in the dataset. Weekly returns were calculated using daily data on share prices (PRCCD).

ROA: The ratio of income before extraordinary items (IB) divided by total assets (AT).

SIGMA: The standard deviation of firm-specific weekly returns for each fiscal year in the dataset. Weekly returns were calculated using daily data on share prices (PRCCD).

SIZE: The natural logarithmic of the market value of equity. Market value of equity is calculated as a stock's annual closing price (PRCC_F) times the amount of common shares (CSHPRI).

Appendix B: Tables

Table A: Industries With The Strongest Industry Effects: CRASH and DUVOL

Table 4 reports the industries with the strongest and weakest effect on stock price crash risk, as determined by a regression where crash risk is regressed on industry dummies and selected control variables (see Appendix A.3. for control variables). The sample is stock data from the Center for Research in Security Prices (CRSP)/Compustat merged database for the period 1985-2022. Crash risk is proxied by a crash dummy (CRASH) in Panel A and B, and by down-to-up volatility for Panel C and D. Crash risk measure definitions can be found in Appendix A.1. The industry dummies are based on the Global Industry Classification Standard (GICS) 6-digit industry classification. The P-value for industry *101010: Energy Equipment & Services* is not available since it is the reference point for the other dummies.

Panel A: Top 10 Industries With Effects of Greatest Magnitude on CRASH Measure, Given $p<0.05$.

GICS Code	Industry Name	Coefficient	P-Value
253020	Diversified Consumer Services	0.148	0
452040	Office Electronics	0.132	0.002
351030	Health Care Technology	0.130	0
402020	Consumer Finance	0.129	0
202020	Professional Services	0.126	0
452010	Communications Equipment	0.125	0
451020	IT Services	0.117	0
402040	Mortgage Real Estate Investment Trusts (REITs)	0.116	0
351020	Health Care Providers & Services	0.115	0
252030	Textiles, Apparel & Luxury Goods	0.114	0

Panel B: Top 10 Industries With Effects of Lowest Magnitude on CRASH Measure, Given $p<0.05$.

GICS Code	Industry Name	Coefficient	P-Value
551010	Electric Utilities	0.028	0.033
601010	Diversified REITs (New Name)	0.033	0.048
501010	Diversified Telecommunication Services	0.034	0.010
203020	Passenger Airlines (New Name)	0.038	0.038
551040	Water Utilities	0.042	0.038
203030	Marine Transportation (New Name)	0.047	0.018
254010	Media (Discontinued 2018)	0.048	0.000
302010	Beverages	0.048	0.003
201010	Aerospace & Defense	0.052	0.000
402030	Capital Markets	0.057	0.000

Panel C: Top 10 Industries With Effects of Greatest Magnitude on DUVOL Measure, Given $p<0.05$.

GICS Code	Industry Name	Coefficient	P-Value
402040	Mortgage Real Estate Investment Trusts (REITs)	0.123	0
601050	Health Care REITs (New)	0.091	0.000
551020	Gas Utilities	0.062	0.000
551040	Water Utilities	0.060	0.004
253020	Diversified Consumer Services	0.048	0.004
601080	Specialized REITs (New)	0.048	0.028
401020	Thrifts & Mortgage Finance (Discontinued)	0.046	0.000
351020	Health Care Providers & Services	0.046	0.000
203040	GroundTransportationNewName	0.043	0.003

Panel D: Top 10 Industries With Effects of Lowest Magnitude on DUVOL Measure, Given $p<0.05$.

GICS Code	Industry Name	Coefficient	P-Value
201050	Industrial Conglomerates	-0.067	0.006
404030	Real Estate Management & Development (Discontinued 2016)	-0.064	0.031
251020	Automobiles	-0.059	0.007
254010	Media (Discontinued 2018)	-0.039	0.002
201040	Electrical Equipment	-0.038	0.003
501020	Wireless Telecommunication Services	-0.036	0.043
201010	Aerospace & Defense	-0.028	0.025
101020	Oil, Gas & Consumable Fuels	-0.022	0.036
401010	Banks	0.021	0.036
202010	Commercial Services & Supplies	0.023	0.035

References

Abernathy, W. J., & Utterback, J. M. (1978). Patterns of industrial innovation. *Technology Review, 80*(7), 40–47.

Agarwal, R., & Gort, M. (1996). The Evolution of Markets and Entry, Exit and Survival of Firms. *The Review of Economics and Statistics, 78*(3), 489.

Ali, A., Klasa, S., & Yeung, E. (2008). The Limitations of Industry Concentration Measures Constructed with Compustat Data: Implications for Finance Research. *Review of Financial Studies, 22*(10), 3839–3871.

Birley, S., & Westhead, P. (1990). Growth and Performance Contrasts between "Types" of Small Firms. *Strategic Management Journal, 11*(7), 535–557.

Botosan, C. A. (1997). Disclosure Level and the Cost of Equity Capital. *The Accounting Review, 72*(3), 323–349.

Callen, J. L., & Fang, X. (2015). Religion and stock price crash risk. *Journal of Financial and Quantitative Analysis, 50*(1-2), 169-195.

Chang, X., Chen, Y., & Zolotoy, L. (2017). Stock Liquidity and Stock Price Crash Risk. *The Journal of Financial and Quantitative Analysis, 52*(4), 1605-1637.

Chen, J.; H. Hong; and J. C. Stein. (2001). Forecasting Crashes: Trading Volume, Past Returns, and Conditional Skewness in Stock Prices. *Journal of Financial Economics, 61*, 345–381.

Clarke, R., Davies, S., & Waterson, M. (1984). The Profitability-Concentration Relation: Market Power or Efficiency? *The Journal of Industrial Economics, 32*(4), 435.

Conrad, J., R. F. Dittmar, and F. Ghysels, 2013, Ex ante skewness and expected stock returns, *The Journal of Finance, 68*, 85–124

Datta, S., Iskandar-Datta, M., & Singh, V. (2013). Product Market Power, Industry Structure, and Corporate Earnings Management. *Journal of Banking & Finance, 37*(8), 3273–3285.

Evans, D. S. (1987). The Relationship Between Firm Growth, Size, and Age: Estimates for 100 Manufacturing Industries. *The Journal of Industrial Economics*, 35(4), 567.

Frankel, R., McNichols, M., & Wilson, G. P. (1995). *Discretionary Disclosure and External Financing. The Accounting Review, 70*(1), 135–150.

Graham, J. R., Harvey, C. R., & Rajgopal, S. (2005). The Economic Implications of Corporate Financial Reporting. *Journal of Accounting and Economics, 40*(1-3), 3–73.

Harvey, C. R., & Siddique, A. (2000). Conditional Skewness in Asset Pricing Tests. *The Journal of Finance, 55*(3), 1263–1295.

Hertzel, M. G., & Officer, M. S. (2012). Industry Contagion in Loan Spreads. *Journal of Financial Economics, 103*(3), 493–506.

Hou, K., & Robinson, D. T. (2006). Industry Concentration and Average Stock Returns. *The Journal of Finance, 61*(4), 1927–1956.

Hutton, A. P.; A. J. Marcus; and H. Tehranian. (2009). Opaque Financial Reports, R^2, and Crash Risk. *Journal of Financial Economics, 94*(1), 67–86.

Janssen, F. (2009). The Conceptualisation of Growth. *The Journal of Entrepreneurship, 18*(1), 21–45.

Jin, L., & S. C. Myers. (2006). R^2 around the World: New Theory and New Tests. *Journal of Financial Economics, 79*(2) , 257–292.

Jovanovic, B., & MacDonald, G. M. (1994). The Life Cycle of a Competitive Industry. *Journal of Political Economy, 102*(2), 322–347.

Karuna, C. (2007). Industry Product Market Competition and Managerial Incentives. *Journal of Accounting and Economics, 43*(2), 275–297.

Kaustia, M., & V. Rantala. (2021). Common Analysts: Method for Defining Peer Firms. *Journal of Financial and Quantitative Analysis.* 56(5), 1505-1536.

Kim, J.-B.; Y. Li; and L. Zhang. (2011a). Corporate Tax Avoidance and Stock Price Crash Risk: Firm-Level Analysis. *Journal of Financial Economics, 100*(3), 639–662.

Kim, J.-B.; Y. Li; and L. Zhang. (2011b). CFOs versus CEOs: Equity Incentives and Crashes. *Journal of Financial Economics, 101*(3), 713–770.

Kothari, S. P., Shu, S., & Wysocki, P. D. (2009). Do Managers Withhold Bad News? *Journal of Accounting Research, 47*(1), 241–276.

Kumar, M. S. (1985). Growth, Acquisition Activity and Firm Size: Evidence from the United Kingdom. *The Journal of Industrial Economics, 33*(3), 327.

Li, J., & Luo, Z. (2020). The impact of product market competition on stock price crash risk. *Journal of Business & Industrial Marketing, 35*(7).

Li, Si, & Xintong Zhan. (2019). Product Market Threats and Stock Crash Risk. *Management Science, 65*(9): 4011-4031.

Lang, L. H. P., & Stulz, R. M. (1992). Contagion and Competitive Intra-Industry Effects of Bankruptcy Announcements. *Journal of Financial Economics, 32*(1), 45–60.

Melicher, R. W., Rush, D. F., & Winn, D. N. (1976). Degree of Industry Concentration and Market Risk-Return Performance. *The Journal of Financial and Quantitative Analysis*, *11*(4), 627.

Porter, M.E. (2008). The Five Competitive Forces That Shape Strategy. *Harvard Business Review, 86*(1), 78–93.

Sapienza, P. (2004). The Effects of Government Ownership on Bank Lending. *Journal of Financial Economics, 72*(2), 357–384.

Schmalensee, R. (1985). Do Markets Differ Much? *The American Economic Review, 75*(3), 341–351.

Slovin, M. B., Sushka, M. E., & Bendeck, Y. M. (1991). The Intra-Industry Effects of Going-Private Transactions. *The Journal of Finance, 46*(4), 1537–1550.

Zhang, X., Tan, J. and Chan, K.C. (2021), Environmental law enforcement as external monitoring: evidence from the impact of an environmental inspection program on firm-level stock price crash risk, *International Review of Economics and Finance, 71*, 21-31.

Yip, G. S. (1982). *Barriers to Entry: A Corporate Strategy Perspective.* Lexington Books.