Financial Distress Risk and Stock Price Crashes

Christoforos K. Andreou[a]
Panayiotis C. Andreou[a,b]
Neophytos Lambertides[a]

December 2020

ABSTRACT

This study uses 462,678 monthly observations of US-listed firms for the period 1990-2018 to document a strong positive relationship between short-term changes in financial distress risk and future stock price crashes. This result is economically significant as a one interquartile increase of the main explanatory variable in any month increases the probability of a stock price crash by 8.33 percent relative to its mean value. The findings withstand controls for a large array of variables, firm-fixed effect estimations, and alternative definitions of distress and crash risk measures; they are also robust to a range of tests conducted to buttress against endogeneity concerns. The study conducts analyses demonstrating that the positive distress-crash risk relationship is driven by managerial opportunism that seeks to camouflage bad news that has an adverse effect on firms' economic fundamentals. Accordingly, the findings corroborate an agency theory explanation for the impact of distress risk on stock price crashes. This study offers practical insights to investors, who should be vigilant of a firm's distress risk, as sudden short-term increases underscore withheld negative information pertinent to crash risk problems.

Keywords: firm-specific stock price crashes; distress risk; bad news hoarding; agency problems; managerial opportunism; financial analysts.

JEL Classification: G12; G19; G32; G33; M40; D89

---

* We would like to thank Lenos Trigeorgis, Tarik Driouchi, Photis Panayides, Alexandros Kontonikas, Jerry Coakley, Agnieszka Trzeciakiewicz, Andreas Savvides, Neil Kellard, Jeffrey L. Callen, Jens Hagendorff, Yitzhak Venezia, Marina Magidou and the seminar participants at the Cyprus University of Technology, Financial Management & Accounting Research Conference (FMARC) 2019, the EFiC Conference 2019, and the International Symposium of Finance 2019 for their useful comments and suggestions. All remaining errors are our own.
[a] Cyprus University of Technology, Department of Commerce, Finance and Shipping, Archbishop Kyprianou 30, Limassol 3603, Cyprus.
[b] Durham University, Durham University Business School, Mill Hill Lane, Durham, DH1 3LB, United Kingdom.
Emails: chk.andreou@cut.ac.cy, panayiotis.andreou@cut.ac.cy, n.lambertides@cut.ac.cy.

1. Introduction

A considerable body of research theorizes that the desire of managers to preserve their wealth and human capital incentivizes them to strategically withhold bad news, which can keep investors' expectations at unjustifiable levels and inflate a firm's stock price beyond its intrinsic value at the expense of shareholders (*e.g.* Jin and Myers, 2006; Bleck and Liu, 2007; Benmelech, Kandel, and Veronesi, 2010). Accordingly, such opportunistic behavior prolongs the false impression investors have regarding the firm's true state of economic fundamentals (Kothari *et al.*, 2009; Hutton, Marcus, and Tehranian, 2009; Kim, Li, and Zhang, 2011a). Keeping the deception up is naturally unsustainable in the long-term and when the volume of negative information becomes overwhelming, managers tend to give up. At that point, the accumulated negative information spills into the market in an abrupt fashion, causing a firm-specific stock price crash.

The burgeoning literature attributes firm-specific stock price crashes to agency-related problems arising from managerial opportunism, which fuels the *bad news hoarding* mechanism (*e.g.* Hutton *et al.*, 2009; Kim *et al.*, 2011a; Callen and Fang, 2013; Andreou *et al.*, 2016; Andreou *et al.*, 2017b). From a different perspective, a number of other studies show that managers of firms facing rising distress risk situations act opportunistically to obfuscate their firm's poor operating performance; for example, by influencing contractual outcomes or misleading stakeholders about their firms' economic fundamentals (DeAngelo, DeAngelo, and Skinner, 1994; Rosner, 2003; Charitou, Lambertides, and Trigeorgis, 2007; Andreou, Lambertides, and Panayides, 2021). Research also suggests that the link between distress risk and managers' career concerns represents one of the reasons why managers persistently withhold bad news (*e.g.* Kothari, Shu, and Wysocki, 2009). Taking these ideas on board, we hypothesize that the negative externalities associated with rising financial distress risk incentivize managers to persistently withhold bad news from investors, a strategy that increases firms' susceptibility to future stock price crashes. Despite the plausibility and research-worthiness of this proposition, to the best of our knowledge, studies have yet to meticulously investigate the relationship between financial distress risk and the future occurrence of stock price crashes. In this respect, our study fills this gap by seeking to empirically discover *a positive distress-crash risk relationship*.

Evidencing a positive distress-crash risk relationship has been an elusive task in the limited number of studies that use financial distress risk in their analyses. Zhu (2016) and Andreou *et al.* (2017b) use firms' financial distress risk—estimated from accounting and market-based models—as just one of a number of control variables in their firm-year panel regression models, and report no notable statistical relationship with future stock price crashes. One reason explaining the absence of a significant relationship relates to the measurement of distress risk using yearly intervals, which is the mainstream approach in corporate finance and accounting studies. Prior evidence supports the view that managers have the capacity to strategically hide bad news pertaining to their firms' true state of economic fundamentals *for up to three or* *more years* without being detected by investors (Jin and Myers, 2006; Bleck and Liu, 2007; Kothari *et al.*, 2009; Hutton *et al.*, 2009; Benmelech *et al.*, 2010; Andreou *et al.*, 2017b). These strategies enable managers to disguise negative information for long periods, keeping the public in the dark regarding its adverse impact on the firm's economic value. Once investors discover a portion of previously withheld negative information that enables them to better discern the firm's (true) state of economic fundamentals, the decline in their expectations will be impounded in the market, quickly driving the firms' distress risk level to higher levels. We argue that such information discovery usually happens in the brief period before managers eventually abandon their efforts to keep the public in the dark. In this vein, a firm's financial distress risk condition, and its implications regarding future crash risk, *cannot be accurately appraised by investors*, when the appraisal occurs much earlier the point in time that managers give up and publicly release stockpiled bad news that triggers a stock price crash. Hence, it may not be possible to witness a positive distress-crash risk relationship when conducting empirical analyses with low frequency data, whereby distress risk is measured at yearly observation intervals and well ahead of the point in time that negative information would start to flow in the market.

We circumvent the above limitation by estimating a firm's distress risk using *monthly observation intervals*. This choice is also motivated by the findings in Chava and Jarrow (2004), who provide evidence that distress risk models estimated with monthly, as opposed to yearly, observation intervals span more timely information and are significantly more accurate in forecasting exercises (see, also, Campbell, Hilscher, and Szilagyi, 2008). We estimate a firm's distress risk level to be the probability to default based on the innovative forecasting KMV-Merton approach, following its possible implementation as in Bharath and Shumway (2008). Studies like those of Hillegeist, Keating, Cram, and Lundstedt (2004) and Vassalou and Xing (2004) substantiate the ability of the Merton (1974) distance to default (*DD*) model to capture timely information about a firm's economic fundamentals much faster than traditional rating models and econometric approaches that rely on accounting data (see, also, Charitou, Dionysiou, Lambertides, and Trigeorgis, 2013). The study by Campbell *et al.* (2008) also supports that models of financial distress exploiting stock market-based variables rapidly incorporate new information about the firm's prospects. Andreou (2015) also shows that short-term changes of distress risk, as measured by the Merton *DD* model, are associated with the higher option-implied moments of future stock returns.

Xerox is a notable real-world example that supports the hypothesized short-term positive distress-crash risk relationship. Its financial distress risk (as captured by the Merton *DD* model in our data) suddenly increased by 23 percent points in November of 2000 compared to the previous month, flagging an early warning for investors regarding the 2000-Q4 results, which turned out to be rather disappointing (largest quarterly loss in a decade). According to the Securities and Exchange Commission, prior to announcing its 2000-Q4 results, and since as early as 1997, Xerox routinely included misleading information in its corporate disclosure.[1] As Xerox management had made a strategy decision to consistently provide obfuscated information to investors regarding the firm's state of fundamentals, Xerox's stock price had become greatly inflated compared to its the intrinsic value, depressing its market-based distress risk situation below true level. By the last quarter of 2000, Xerox's management was unable to continue manipulating its financial statements and was forced to release the bad news in the market, causing a severe stock price decline of about 30 percent in December of 2000.

The Xerox case can provide anecdotal support for the existence of a positive distress-crash risk relationship *that spans in the short-term*. Interestingly, Xerox's sudden increase in distress risk level between October and November of 2000 (observed one month ahead of the crash incidence), happened because negative information had previously been withheld from the market, which when discovered it quickly worsened the investors' expectations regarding the firms' true economic value. This stylized fact resonates with studies supporting that releasing a portion of (hitherto private or withheld) negative information is likely to spill into the market in a gradual fashion, ahead of other important events or announcements. For instance, Hong and Stein (1999) discuss the notion of gradual information diffusion, whereby as a result of either the technology of information distribution, or investor segmentation and specialization, certain pieces of value-relevant, privately-held information will arrive in the hands of some investors before others. Hong and Stein (2003) also discuss an information structure process whereby managers have the tendency to release negative firm-specific information in a piecemeal fashion. Further, Roychowdhury and Sletten (2012) elaborate on the notion that the extent of information made available in the market is most influenced by market participant scrutiny, such as financial analysts who monitor firm performance and produce high-quality information, rather than voluntary disclosures by managers.[2]

The Jin and Myers (2006) model is the most widely admitted paradigm in crash risk literature regarding information structure dynamics. According to this model, the withheld negative information spills into the market *abruptly and all at once* at the point where managers give up (*i.e.* become unwilling or unable) to continue concealing it. Nevertheless, based on the arguments above (*e.g.* Hong and Stein, 1999, 2003; Roychowdhury and Sletten, 2012; Callen and Fang, 2015; An *et al.*, 2020; Deng *et al.*, 2020), it is reasonable to assume that at least a portion of the hitherto undisclosed bad news that managers are strategically concealing from the market spills into the market in the short period *preceding* the aforementioned tipping point. In this respect, the discovery of such negative information increases the

[1] More information is provided here: https://www.sec.gov/litigation/complaints/complr17465.htm.

[2] An, Chen, Naiker and Wang (2020) argue that investors can anticipate firms' future bad news by gleaning information revealed by media scrutiny, which can expose managerial misconduct and other problems even in the absence of firms' own disclosures. Studies like those of Callen and Fang (2015) and Deng, Gao, and Kim (2020), add that short sellers analyze information from multiple information channels (beyond earnings as disclosed in financial statements) to promptly detect bad news hoarding activities by firms; they then short those firms' stocks in anticipation of a price crash.

firms' distress risk level within a short period of time, as investors start revising their expectations downwards regarding the firms' true state of economic fundamentals. Our analyses therefore focus on the *short-term changes* in distress risk and investigate their implications on the future occurrence of stock price crashes.[3]

Using a large sample featuring 462,678 monthly observations of US-listed firms for the period 1990-2018, and consistent with our prediction, this study documents a strong positive relationship between the 3-month changes in distress risk (measured in month *t*-1) and future incidence of stock price crashes (measured in month *t*). The economic significance of short-term changes in distress risk is 8.33 percent of the crash sample mean value, and much bigger than that of other prominent crash risk determinants (*e.g.* opacity is 3.01 percent, investor heterogeneity is 5.77 percent, *et cetera*). This distress-crash risk relationship remains significantly positive and economically meaningful, even after controlling for a wide array of covariates as identified by extant studies (*e.g.* Chen *et al.*, 2001; Hutton *et al.*, 2009; Kim *et al.*, 2011a, 2011b; Andreou, 2015; Kim and Zhang, 2016; Andreou *et al.*, 2017b). The results are strong and robust when using another financial distress measure, or when using five alternative crash risk measures, even when considering the joint inclusion of the variables used to calculate the distress risk measure.

We conduct a battery of identification strategies to address potential endogeneity. We use a lead-lagged design as a first step to mitigate any reverse causality concerns. Interestingly, the short-term changes in a firm's financial distress risk measured in month *t*-2, or *t*-3, or even *t*-4 continue to exhibit a strong significant and positive association with the incidence of a stock price crash in month *t*. We conduct two additional tests, whereby we use stock price crashes as the main explanatory variable and future changes in distress risk as the dependent variable, as well as performing regression estimations by dropping the firm-month observations that could induce the reverse relationship. Both tests help us to reject the reverse causation explanation for our main findings. The positive distress-crash risk relationship remains strong when we use firm-fixed effects regression to control for unobserved time-invariant firm-specific characteristics. Additionally, we continue to observe the strong positive distress-crash risk relation when using a two-stage instrumental variable estimation and a quasi-experimental setting that exploits the enactment of the Sarbanes-Oxley Act as an exogenous shock.

In supplementary analyses, we show that bad news withholding is the underlying reason that substantiates the positive link between short-term changes in distress risk and future stock price crashes. First, we provide results showing that short-term changes in distress risk associate with situations whereby

[3] Our reliance on short-term changes in distress risk underscores our intent to measure the cumulative effect of bad-news flows in the market (within that short period of time) and how informative the news is regarding the firm's state of economic fundamentals. Our approach is similar in spirit to Roychowdhury and Sletten (2012), who use cumulative (buy-and-hold) returns between fiscal quarters as a comprehensive measure of news released within that 3-month period, in order to investigate earnings informativeness in bad-news quarters relative to good-news quarters.

managers try to disguise bad news through earnings management manipulation. Second, in cross-sectional analyses, we find that the predictive power of short-term changes in distress risk with respect to future crashes is stronger in environments where investors are faced with higher information asymmetries and it is therefore easier for managers to withhold bad news and accumulate adverse information. Third, we rely on analysts' buy-sell revisions to support the working hypothesis, whereby the increases of short-term changes in distress risk capture negative information that gradually spills into the market ahead of the crash event.

In terms of contributions, our study is the first to thoroughly examine the impact of financial distress risk on the future occurrence of stock price crashes. Most crash risk studies to date use accounting-based leverage as a control variable in their baseline regression models. Admittedly, accounting-based leverage is a poor proxy for a firm's distress risk as assessed by the market. For example, Hillegeist *et al.* (2004, p. 6) claim that: "*[market-based distress risk] estimates are statements about the likelihood of future events, [while] the financial statements are designed to measure past performance and may not be very informative about the future status of the firm*". It is not surprising then that many studies (*e.g.* Hutton *et al.*, 2009; Kim *et al.*, 2011a; Kim *et al.*, 2016; Chang *et al.*, 2017; Deng *et al.*, 2020) have not found statistical support for the positive distress-crash risk relation when leverage is used as a proxy. Interestingly, many other studies (*e.g.* Andreou *et al.*, 2016; Callen and Fang, 2017; Chen *et al.*, 2017; Li and Zeng, 2019; An *et al.*, 2020) report significantly *negative* coefficients for financial leverage, which is contrary to the expected relationship. Accordingly, our study seeks to clear up inconclusive inferences in prior studies and unveil a positive distress-crash risk relationship.

Related to the above, Zhu (2016) and Andreou *et al.* (2017b) use firms' financial distress risk as a control variable in their *firm-year* panel regression models. Again, they report no notable statistical relationship with future stock price crashes. Evidently, most crash risk studies focus on estimations using yearly measurement intervals.[4] To the best of our knowledge, our study is also the first to derive results from *firm-month* panel regression models. In this respect, our research design enables us to better assess the crash risk-related implications of financial distress risk. At the same time, our methods lay the foundation for future research on crash risk aiming to discover the short-term dynamics of various crash risk antecedents.

The predominant view in extant crash risk studies follows the agency-based paradigm of Jin and Myers (2006) predicating that bad news is withheld and piles up until it reaches a certain tipping point; what follows is a *sudden and complete* release of accumulated negative information into the market, immediately causing a large and abrupt firm-specific stock price crash. In this vein, our findings support the bad news

[4] A handful of exceptions—Chen *et al*. (2001) and Ak, Rossi, Sloan, and Tracy (2016)—consider semi-annual measurement intervals, while Callen and Fang (2015), Chen, Kim, and Yao (2017) and Ni, Peng, Yin, and Zhang (2020) consider quarterly measurement intervals.

hoarding mechanism as delineated in this paradigm, and they corroborate the strand of studies that elaborate on financial distress conditions that induce managers to behave opportunistically (*e.g.* Charitou *et al*., 2007; Kothari *et al.*, 2009; Chava, Kumar, and Warga, 2010, *et cetera*).

At the same time, the Jin and Myers (2006) paradigm precludes the possibility that certain investors may discover some portion of the withheld negative information and impound it in the stock market ahead of the crash event. However, our findings suggest that a firm's distress risk increases in the brief period that precedes the crash incidence because some portion of the withheld bad news spills into the market ahead of the crash event. This evidence challenges the information structure of Jin and Myers' paradigm, lending credence to the notion that negative firm-specific information spills in a piecemeal fashion ahead of the crash event, something that supports the arguments in the studies of Hong and Stein (1999, 2003) and Chen *et al.* (2001). In this way, our findings inform an emerging strand of crash risk literature (*e.g.* Callen and Fang, 2015; An *et al.*, 2020), suggesting that investors use information channels other than management-initiated disclosure to discover withheld bad news that adversely affect a firm's true state of economic fundamentals.

The remainder of this study proceeds as follows: Section 2 describes the data, measurements, and the methodology; Section 3 presents the summary statistics and univariate analysis; Section 4 provides empirical analyses of the distress-crash risk relationship, as well as a robustness analysis; and Section 5 presents additional analyses. Finally, Section 6 discusses the practical aspects of our findings and provides a conclusion to the study.

2. Data, variable measurement, and regression models

2.1. Sample data

Our sample includes US-listed firms in the NYSE, AMEX, and NASDAQ with data available in the Compustat Quarterly and Center for Research in Security Prices (CRSP) databases, excluding financial services (SIC 6000-6999) and utilities firms (SIC 4900-4999). To preclude look-ahead bias, we match market data with quarterly accounting data by lagging them by three months. Additionally, we exclude: (i) firm-year observations with an average market capitalization below 30 million dollars to minimize the influence of small stocks;[5] (ii) firm-month observations with fewer than 26 weekly returns within the

[5] An average market capitalization of 30 million dollars during the fiscal year corresponds to stocks that have an average price of 2.50 dollars, a condition used by many other studies (*e.g.* Andreou *et al.*, 2017b). Conditioning on average market capitalization is preferable since, quite frequently, low priced stocks tend to have high market capitalizations. For instance, using CRSP data from 2018, over 170 stocks have a price below 2.50 dollars, while their market capitalization ranges between 0.5 and 4.7 billion dollars. All results are similar if we instead exclude observations that have a price below 2.50 dollars.

estimation period to minimize the influence from illiquid trading on the estimation of crash risk; (iii) firms that went bankrupt as documented in the UCLA-LoPucki Bankruptcy Research Database to eliminate the possibility that findings are driven by the special characteristics of such firms; and (iv) observations with insufficient financial data to calculate the main variables of our analysis. The final sample covers the period from January 1990 to December 2018 and consists of 462,678 *firm-month* observations that correspond to 4,855 unique firms.

2.2. Measuring stock price crashes

We employ six different month-based stock price crash measures, particularly, three dichotomous and three continuous operationalizations. The primary measure used to conduct our analyses—*CRASH*—is an indicator variable set equal to one for months when a firm experiences an *extreme firm-specific left-tail outcome*, and zero otherwise. We have a preference for this dichotomous definition because: (i) it aligns with the theoretical underpinnings in Jin and Myers (2006), delineating a stock price crash as being an *idiosyncratic, large negative outlier* in the distribution of returns, and (ii) it has been widely adopted by researchers in the ambit of empirical crash risk studies (*inter alia,* Hutton *et al.*, 2009; Kim *et al.*, 2011a; Kim *et al.*, 2016; Zhu, 2016; Andreou *et al.*, 2017a, Andreou *et al.*, 2017b; Chang *et al.*, 2017).

We first compute firm-specific weekly returns using the following expanded index model:

$$r_{i,w}=\alpha_i+\beta_{1,i}r_{m,w-2}+\beta_{2,i}r_{m,w-1}+\beta_{3,i}r_{m,w}+\beta_{4,i}r_{m,w+1}+\beta_{5,i}r_{m,w+2}+\varepsilon_{i,w}\ , \quad (1)$$

where $r_{i,w}$ is the return of firm *i* in week *w* and $r_{m,w}$ is the CRSP value-weighted market index return in week *w*. Following the crash risk literature, the *firm-specific weekly return* for firm *i* in week *w*, namely $R_{i,w}$, is defined as the natural logarithm of one plus the residual return of *Eq.* (1):

$$R_{i,w}= ln\left[1+\varepsilon_{i,w}\right]. \quad (2)$$

Accordingly, *CRASH* is an indicator variable set equal to one for month *t* if within this month the firm's stock experiences at least one firm-specific weekly return that falls more than 3.09 standard deviations below the mean firm-specific weekly return over the estimation period, with 3.09 chosen to generate a frequency of 0.1 percent in the normal distribution; otherwise, it is set equal to zero. To define *CRASH* in month *t*, the model in *Eq.* (1) is estimated with a rolling regression approach using the most recent 52 calendar weeks, whereby month *t* is included as the last calendar month in the estimation data. In this fashion, *CRASH* potentially captures the occurrences of stock price crashes in a timelier fashion and fits well with the needs of our study that conducts its analyses using monthly measurement intervals.

We test the robustness of our baseline results with two more dichotomous variables. Following Kim *et al.* (2011a) and Andreou *et al.* (2017b), likewise we define *CRASH_3.2* by using 3.20 standard deviations as the threshold point. Additionally, we define *CRASH_20PRC* to be equal to one for month *t* if within this month the firm experiences at least one market-adjusted weekly return (*i.e.* stock weekly return minus CRSP value-weighted market index weekly return) that falls more than -20 percent, and zero otherwise. Unlike the other two dichotomous measures, *CRASH_20PRC* is a purely model-free estimate since its computation does not rely on the index model as per *Eq*. (1). More importantly, it does not rely on past stock returns' data since it is computed by using information from month *t* (whereas the change in distress risk uses information up to month *t*-1). Ergo, *CRASH_20PRC* is estimated with information that is completely disjoint from the information used to estimate the main explanatory variable.

We also provide robustness results using other widely applied continuous stock price crash risk measures (*e.g.* Chen *et al.*, 2001; Hutton *et al.*, 2009; Kim *et al.*, 2011a). Such measures intend to capture the negative asymmetry in a firm's stock returns distribution, in essence by capturing stocks that are merely more "crash prone", that is, subject to more left-skewed distribution. However, negative asymmetry in returns is possible to arise by the presence of several *less extreme* negative returns, something that does not necessarily comply with the notion that stock price crash risk represents the likelihood of an extreme negative firm-specific return outlier (see, also, Ak *et al.*, 2016; Andreou *et al.*, 2017a).[6] Against this backdrop, we complement our analyses by providing robustness tests using three continuous crash risk measures capturing information about the conditional skewness of the return distribution.

The first continuous measure is the negative coefficient of minimum return, *NCMRET*. It is defined in the spirit of Ak *et al*. (2016) as:

$$NCMRET_{i,t}=\frac{-min(q_{i,w},q_{i,w-1,\ldots,}q_{i,w-n+1,})}{\sqrt{\frac{\sum MAR_i^2}{(n-1)}}}, \tag{3}$$

where $q_{i,w}$ is the firm's *i* market-adjusted return in week *w* (stock return, $r_{i,w}$, minus CRSP value-weighted return, $r_{m,w}$) over the most recent *n* weeks including the weeks falling in month *t.* For estimation purposes *n* equals 26. The denominator of *Eq*. (3) features the standard deviation of market-adjusted returns, whereby MAR_i captures the sequence of mean squared differences covering the 26-week period before calendar week *w*-*n*. The minus sign is used so that a higher value of *NCMRET* (same for *NCSKEW* discussed shortly) indicates a more left-skewed distribution of firm-specific weekly returns, that is, a higher level of stock price crash risk.

[6] Negative asymmetries in stock returns can arise, *inter alia*, due to the exogenous stochastic process generating information relating to divergence in opinion among investors (*e.g.* Hong and Stein, 2003) and volatility feedback effects (*e.g.* French, Schwert, and Stambaugh, 1987), which rely on theoretical arguments that diverge from the bad news hoarding mechanism as in Jin and Myers (2006). This regularity can rationalize why the correlation between the dichotomous *vs.* continuous crash measures reported by prior studies is not particularly high and falls significantly below unity. For instance, the Pearson correlation coefficient between *CRASH* and *NCSKEW* is 0.490 in Kim and Zhang (2016), 0.63 in Kim *et al.* (2016), 0.51 in Andreou *et al.* (2017a), and 0.62 in Chang *et al.* (2017).

The second continuous measure is the negative coefficient of skewness, *NCSKEW*. It is defined as the third moment of firm-specific weekly returns divided by the standard deviation of firm-specific weekly returns raised to the third power. Specifically, it is calculated as:

$$NCSKEW_{i,t} = \frac{-\left[n(n-1)^{\frac{3}{2}} \sum R_{i,w}^3\right]}{(n-1)(n-2)\left(\sum R_{i,w}^3\right)^{\frac{3}{2}}}, \tag{4}$$

where $R_{i,w}$ is the firm-specific weekly return for firm *i* in week *w* following *Eq*. (2), and *n* is the number of firm-specific weekly returns included in its estimations. To compute *NCSKEW* in month *t*, the firm-specific weekly returns as per *Eq*. (2) are estimated with a rolling regression that utilizes the most recent 52 calendar weeks, inclusive of the weeks that fall within month *t* (likewise for *DUVOL* below).

The third continuous crash risk measure is the down-to-up volatility, *DUVOL*. It is defined as the logarithm of the standard deviation of "down" weeks over the standard deviation of the "up" weeks. A "down" ("up") week is when the firm-specific weekly return ($R_{i,w}$) is below (above) the estimation period's mean weekly return. Specifically, it is calculated as:

$$DUVOL_{i,t} = log\left(\frac{(n_u-1)\sum_{DOWN} R_{i,w}^2}{(n_d-1)\sum_{UP} R_{i,w}^2}\right), \tag{5}$$

where n_u and n_d are the number of "up" and "down" weeks that fall in the most recent 52 calendar weeks.

2.3. Measuring financial distress risk

We proxy for financial distress risk using the firm's specific probability to default as computed by Merton *DD* model. We employ the "naïve" Merton *DD* model as delineated in Bharath and Shumway (2008), in which the inputs are either inferred using monthly stock-based data or are observable from the firm's quarterly financial statements. The authors argue that the naïve approach improves the accuracy of predictions and avoids tackling any equations or estimating any difficult quantities in its construction in the course of making valid short-term, out-of-sample forecasting inferences (see, also, Charitou *et al.*, 2013; Andreou, 2015). Accordingly, a firm's probability of default at the debt's maturity is computed monthly as follows:

$$DD_{i,t} = \frac{ln\left(\frac{V}{D}\right) + \left(AR_{i,t-1} - 0.5\sigma_{BS}^2\right) T}{\sigma_{BS}\sqrt{T}}, \tag{6}$$

where $DD_{i,t}$ is the distance to default for firm *i* in month *t,* and *V* is the total value of the firm's assets in month *t,* which equals the firm's market value of equity (*ME*) plus the face value of debt (*D*) in month *t*. $AR_{i,t-1}$ represents the expected return on the firm's total asset value for firm *i* measured in month *t*-1 calculated by cumulating the past year's monthly returns. The volatility of the firm's total asset value returns (σ_{BS}) for month *t* is calculated as the weighted average of the volatilities of a firm's equity and debt. Further,

T is the assumed firm's debt maturity in month t, which is always set equal to one year. Detailed variable definitions are described in the Appendix.

Subsequently, distress risk (*DR*) for firm i in month t is the probability to default, calculated as the cumulative standard normal distribution of the negative distance to default, hence:

$$DR_{i,t}=N(-DD_{i,t}), \tag{7}$$

Thereupon, and in accordance to our research scope aiming to explore the distress-crash risk relationship in the short-term, our analyses utilize a 3-month *change of distress risk* as follows:

$$\Delta DR_{i,t-1}=DR_{i,t-1} - DR_{i,t-4} \tag{8}$$

The short-term changes in distress risk as per *Eq.* (8) feature firm-specific information stemming from both quarterly accounting and monthly market-based variables, and, hence, they are perceived to capture the market-based estimates of investors' expectations regarding a firm's state of economic fundamentals for the near future. As such, we presume that they capture critical information in a timely fashion, underscoring the firm's prospects, as also postulated by prior literature (Vassalou and Xing, 2004; Hillegeist *et al.*, 2004; Bharath and Shumway, 2008; Campbell *et al.*, 2008; Andreou, 2015).

This definition is a sensible operationalization to capture short-term variations in distress risk that are driven by sudden (*e.g.* negative) revisions in investors' expectations regarding a firm's state of economic fundamentals. The latter conjecture aligns with Roychowdhury and Sletten (2012), who rely on cumulative stock returns falling between two consecutive fiscal quarters to empirically measure the flow of news in the market during that 3-month period. As they claim, returns are a comprehensive measure of news released by all available information channels. By the same token, short-term changes in distress risk should comprehensively capture the flow of news coming from different sources and provide concise information on a firm's true state economic fundamentals. As evinced by the estimation of *DR* in *Eq.* (7), in addition to the informativeness spanned by past cumulative stock returns, the computation of Merton's *DD* considers additional information emanating both from accounting and market-based variables. For example, *inter alia*, the model accounts for a firm's market-based asset volatility, which is a crucial variable because it captures the likelihood that the value of the firm's assets will decline to such an extent that the firm will be unable to repay its debts (Hillegeist *et al*., 2004).

2.4. Control variables

We rely on several control variables relevant to this context calculated at the monthly frequency (detailed descriptions in the Appendix). We use nine baseline control variables that the literature identifies as important covariates of distress and crash risk. Following Chen *et al.* (2001) and Hutton *et al.* (2009), we control, among other variables, for: firm size (*SIZE*), market-to-book ratio (M/B), financial leverage (*LEV*) and return on assets (*ROA*). He and Ren (2017) report that financial constraints increase future stock price crash risk. Hence, at the end of each month, we likewise measure firms' financial constraints by the Size-Age (*SA*) index as developed by Hadlock and Pierce (2010). Controlling for this characteristic enables us to show that the predictive power of short-term changes in distress risk on the future occurrence of stock price crashes far exceeds any explanatory power associated with the firm's financial constraints.

We control for past momentum in monthly stock returns because, as Chen *et al.* (2001) discuss, stock returns have time-varying influences on stock skewness and stocks with high past returns tend to be more crash prone. Accordingly, we use the cumulative stock returns from month *t-4* to month *t-1* (*RET*). This variable enables us to offset any concern about the potential mechanical impact of *ΔDR* on future stock price crashes coming from recent stock market swings on the equity value, which previous literature has shown to be an important determinant of distress risk (*e.g.* Campbell *et al.*, 2008; Hillegeist *et al.*, 2004). Further, Roychowdhury and Sletten (2012) suggest that recent cumulative stock returns capture the effect of news that are impounded in the market. We therefore include *RET* in all our regression models, enabling us to alleviate concerns that *ΔDR* merely reflects past stock return informativeness.

Hong and Stein (2003) show that stock price crashes are more pronounced around periods of heavy trading volume that most likely feature differences in investor opinions. Similarly, Chen *et al.* (2001) demonstrate that firms that have experienced an increase in trading volume relative to trend over the past six months encounter more crashes. Thus, we control for differences in investors opinions by using the proxy of Chen *et al.* (2001), namely, detrended stock trading volume (*DTURN*).

We also control for financial reporting opacity (*OPACITY*) based on an indicator of earnings management following Hutton *et al.* (2009), who find a positive relationship between opacity and future stock price crashes. Finally, we control for the negative skewness of the prior year's firm-specific stock returns (*NCSKEW*), since previous studies indicate that firms with high past negative skewness are more prone to future stock price crashes (*e.g.* Kim and Zhang, 2016; Andreou *et al.*, 2017a).

2.5. Regression model specification

The main empirical investigations are conducted with a logit regression model specification as follows:

$$CRASH_{i,t} = \alpha_0 + a_1 \Delta DR_{i,t-1} + \sum_{k=2}^{K} a_k CONTROLS_{i,t-1} + e_{i,t}, \tag{9}$$

where $CRASH_{i,t}$ is our primary dependent variable that takes the value of one when there is a crash occurring in month *t*, and zero otherwise. *CONTROLS* is the vector that includes the baseline control variables (namely, *SIZE, M/B, LEV, ROA, SA, RET, DTURN, OPACITY, NCSKEW*), whereby all are measured in month *t*-1 (or at a more distant point in the past). In principle, we rely on a *lead-lagged relationship* to safeguard our analyses from potential simultaneous causality problems. Subsequently, we provide robustness evidence, showing a statistically positive coefficient for *ΔDR* even when measured up to five periods before the measurement of *CRASH*.

The estimation of the logit regression model always includes industry and fiscal year fixed effects (unless otherwise specified). For industry effects, we use the 48-industry classifications by Fama and French (1997). Furthermore, standard errors are always clustered at the firm level. All continuous variables included in our analyses are winsorized at the 1st and 99th percentiles to mitigate the effect of outliers. To put the variables on the same scale, all continuous variables are standardized to have a mean value of zero and a standard deviation of one.

3. Summary statistics and univariate analyses

Table 1, Panel A shows yearly summary statistics for our main dependent variable (*CRASH*). Although all subsequent results are based on *firm-month* analyses, we first present *firm-year* stock price crash statistics for comparability with the plethora of prior crash risk studies that employ analyses using yearly measurement intervals. For the period 1990–2018, our sample includes in total 42,837 firm-year observations [column (2)], of which 9,010 firm-years [column (3)], or 21.03 percent of the total sample [column (4)], feature stock price crashes. As expected, the mean annualized return for firm-year observations *with* a stock price crash is substantial and equals -17.56 percent [column (5)], whilst the mean annualized return for firm-year observations *without* a stock price crash is 1.91 percent [column (6)]. Both the prevalence and the magnitude of the crashes in our sample are largely consistent with the statistics reported by prior studies (*e.g.* Hutton *et al.*, 2009; Kim and Zhang 2016; Andreou *et al.*, 2017b). Significantly, the mean value of the 3-month changes in distress risk across years, as measured one month before each stock price crash event, is 1.07 percent [column (7)]. The mean value of the 3-month changes in distress risk for firms that did not crash is instead -0.23 percent [column (8)]. The difference in the mean values of the 3-month changes in distress risk between *crash firm-years* [column (7)] *vs. non-crash firm-years* [column (8)] is equal to 1.30 percent [column (9)] and highly statistically significant (p-value < 0.01). This indicates that firms that had an incidence of stock price crash experienced a notable increase in their distress risk in the short period preceding the crash event.

Table 1, Panel B exhibits the anatomy of monthly stock price crashes for the four quarterly earnings announcement (QEA) months compared to the remaining eight (non-QEA) months. Out of the 462,678 firm-month observations, a stock price crash occurs in 9,988 of them. The discrepancy in frequency between the monthly (9,988) *vs.* yearly (9,010) crashes is expected because it is plausible for a firm to experience two or more monthly crashes within a fiscal year. In total, 5,416 (or 54 percent) of these monthly crashes occur during QEA months and 4,572 (or 46 percent) during non-QEA months. This evidence suggests that crashes are not exclusively triggered in periods when managers release formal disclosure (*e.g.* quarterly financial statements) to the public, but in fact quite commonly occur in periods when information is likely diffused in the market through other channels (*e.g.* through analyst forecasts and revisions, short-sellers' scrutiny, media commentary, fiduciary agent scrutiny) as delineated in certain studies (Roychowdhury and Sletten, 2012; Callen and Fang, 2015; An *et al.*, 2020, *et cetera*).

Table 1, Panel C primarily analyzes the ΔDR behavior by QEA and non-QEA months across crash *vs.* non-crash months. Two noteworthy observations emerge. First, by analyzing crashes across the four QEA months, we observe that their occurrence increases monotonically across QEA 1 (1,213) and QEA 2 (1,393), reaching its peak in the third month of QEA 3 (1,738), and then exhibits a decline in the month of QEA 4 (1,072). This empirical regularity in our data resembles a situation recorded by Roychowdhury and Sletten (2012), who estimate earnings' informativeness with respect to bad news by fiscal quarter and discuss that it increases across quarters, reaching its peak in the third fiscal quarter, and then exhibits a slight decline in the fourth.[7] Second, similar to the evidence observed with the yearly crash analysis in Panel A, ΔDR_{t-1} is statistically higher in the period before a crash month (for both QEA and non-QEA months) [column (4)] *vis-à-vis* the periods of non-crash months [column (5)]. Interestingly, ΔDR_{t-1} is particularly heightened in the 3-month period that precedes a stock price crash incidence, which is evidence for a positive distress-crash risk relationship. Additionally, the highest value of ΔDR_{t-1} within the crash months [column (4)]—observed for the non-QEA months—equals 1.89 percent (p-value < 0.01), compared to 0.81 percent (p-value < 0.01) that is observed for the QEA months, with the mean difference between these two cases being equal to 1.08 percent (p-value < 0.01). The latter implies more (severe) negative information flow in non-earnings announcement months. The evidence also suggests that investors, in their effort to discover (hoarded) negative information and infer its implications regarding the firm's prospects (*e.g.* possibility for an imminent adverse firm performance), probably collect and analyze information from a variety of channels other than the financial statements released during the quarterly earnings announcement months.

[Insert Table 1, here]

Table 2 reports summary statistics and Table 3, the Pearson's correlation coefficients. Accordingly, the primary explanatory variable, ΔDR, has a mean of 0 percent and a standard deviation of 9.4 percent. The average value of $SIZE_{t-1}$ and M/B_{t-1} is 6.614 and 3.167 respectively, which is consistent with prior studies (*e.g.* Kim *et al.*, 2011a, 2011b; Andreou *et al.*, 2017b; An *et al.*, 2020). Mean $DTURN_{t-1}$ is 0.001 with a

---

[7] Roychowdhury and Sletten (2012, p. 1699) interpret this result as follows: as the year-end audit approaches, managers release more bad news via the earnings reporting process, causing the observed rise in earnings' informativeness over the first three fiscal quarters. In the fourth quarter, the imminent release of audited annual results can generate intensified scrutiny of the firm by market participants, even prior to the earnings announcements, prompting managers to provide more disclosures.

standard deviation of 0.078, which is in line with other studies (Chen *et al.*, 2001; Callen and Fang, 2015). The mean values of $OPACITY_{t-1}$ and ROA_{t-1} are 0.550 and 0.008, respectively. In general, the distributional characteristics of the variables are qualitatively similar to those of previous studies (Chen *et al.*, 2001; Hutton *et al.*, 2009; Kim *et al.*, 2011a, 2011b; Callen and Fang, 2013; Andreou *et al.*, 2016; Chang, Chen, and Zolotoy, 2017), despite the different time-frequency (monthly *vs.* annual data) and time period of our study.

[Insert Table 2, here]

Table 3, Panel A shows that the correlation between the 3-month changes in distress risk (ΔDR_{t-1}) and the main crash risk measure ($CRASH_t$) is positive (coefficient = 0.020) and statistically significant (p-value < 0.01), providing univariate evidence for *a positive distress-crash risk relationship*. In absolute terms, the correlation of ΔDR_{t-1} with $CRASH_t$ ranks second and only slightly behind that of RET_{t-1}, which equals -0.024. Gauging (in absolute terms) the correlations of $CRASH_t$ with other eminent determinants of crash risk, like for example $OPACITY_{t-1}$ (coefficient = 0.001), ROA_{t-1} (coefficient = 0.007) and $DTURN_{t-1}$ (coefficient = 0.012), reveals a notable level of strength between ΔDR_{t-1} and $CRASH_t$.[8] The correlations of ΔDR with other main control variables are generally rather weak, the only exception being the one with *RET* (coefficient = -0.298), which is expected since a firm's past returns enter in the computation of Merton's *DD* as per *Eq*. (6). This observation also justifies our choice of including *RET* in all regression models to enable us to identify the incremental impact of ΔDR. Notably, the correlation of SA_{t-1} and $CRASH_t$ is -0.011 (p-value < 0.05), whilst the correlation between ΔDR_{t-1} and SA_{t-1} is only -0.003. The latter suggests that the two quantities seem to be orthogonal and potentially feature two distinctive, albeit related, channels through which managers accumulate negative information and increase their firms' susceptibility to future crashes. Other correlation coefficients reported in this table have in general the expected sign and exhibit rather low values.

Table 3, Panel B shows the correlation coefficients between the crash risk measures. As expected, the correlation coefficients between the dichotomous stock price crash measures are rather high. The correlation coefficients between the dichotomous and the other three continuous crash measures are lower and range between 0.112 and 0.310 (p-values < 0.01). Again, this pattern is expected because the dichotomous measures capture extreme negative returns featuring crash incidences in the spirit of Jin and Myers (2006), whilst the continuous measures capture the conditional skewness of the return distribution in the spirit of Chen *et al.* (2001).

---

[8] The magnitude of the correlation coefficient between our main explanatory variable (ΔDR) and future crash risk is comparable to those reported in recent studies that employ large-scale data sets. For instance, Chen *et al.* (2017), who employ 157,722 firm-quarter observations, report correlation coefficients ranging from 0.029 to 0.036 (in absolute terms) between their explanatory variable (*i.e.* earnings smoothing) and 1-quarter-ahead crash risk measures.

[Insert Table 3, here]

4. Empirical analyses of the distress-crash risk relationship

4.1 Baseline and expanded regression models

We move forward with multivariate regression analyses to investigate the relationship between the 3-month changes in distress risk and the likelihood of stock price crashes. Table 4 presents results in the spirit of the logit regression model as described by *Eq.* (9).

Model (1) presents the relationship between ΔDR_{t-1} and $CRASH_t$, without including any of the baseline control variables in the specification. In accordance with our expectations, the coefficient of ΔDR_{t-1} is positive and highly statistically significant (p-value < 0.01). Model (2) includes the vector of our baseline controls. Despite the inclusion of these nine control variables, *ΔDR* remains highly statistically positive with a coefficient value of 0.086 (p-value < 0.01). The strength of *ΔDR* in explaining the likelihood of a future stock price crash compares very favorably with other important and widely recognized crash risk determinants (Chen *et al.*, 2001; Hutton *et al.*, 2009; Kim et al., 2011a); for example, the coefficient of *ΔDR* is 1.46 times bigger than that of *DTURN*, 1.51 times bigger than that of *ROA*, and 2.77 bigger than that of *OPACITY*. As reported in column (3), *ΔDR*'s economic impact on the likelihood of a stock price crash is high and stands at 8.33 percent.[9] The economic impact of *ΔDR* on the likelihood of a future stock price crash is third in order, and greatly exceeds the economic significance of the previously mentioned prominent crash risk predictors; for example, the economic significance of *OPACITY* is 3.01 percent, of *DTURN*, 5.77 percent and of *ROA*, 5.53 percent. These comparisons qualify distress risk as a highly influential—and as yet unexplored—antecedent of crash risk.

With respect to the control variables and contemplating prior evidence (*e.g.* Chen *et al.*, 2001; Zhu, 2016), the coefficients for both *SIZE* and *DTURN* are positive and highly statistically significant (p-value < 0.01), suggesting that the incidence of stock crashes increases with the firm's size and investors' heterogeneity. Moreover, the coefficient of *M/B* is positive while statistically not significant, that of *ROA* is significantly positive (p-value < 0.01), and finally the coefficient of *LEV* positive and marginally statistically significant at the 10 percent level. Like in Kim *et al.* (2016), we find that the returns (*RET*) are negative related to future stock price crashes (p-value < 0.01). We also find the positive association between reporting *OPACITY* (p-value < 0.01) and crash risk first documented by Hutton *et al.* (2009), and a strong positive coefficient (p-value < 0.01) for the one-year-lagged value of *NCSKEW,* consistent with prior

[9] The marginal effects are computed based on Hutton *et al.* (2009), by comparing crash risk at the 25th and 75th percentile values of each variable, while holding all other variables at their mean values. Subsequently, each variable's economic significance is computed by dividing the marginal effect by the mean value of unconditional probability of stock price crash in the sample.

evidence (*e.g.* Andreou *et al.*, 2017a). Finally, the results show that the positive distress-crash risk relationship strongly withstands the inclusion of *SA*.

[Insert Table 4, here]

The remaining models in Table 4 provide various tests to investigate the robustness of the positive distress-crash risk relationship. Models (3) and (4) re-estimate the baseline model specification of model (2) but with a modification in the dependent variables. As evinced in Table 1, Panel B, crashes occur with almost the same frequency across QEA (54 percent) and non-QEA (46 percent) months, indicating that a considerable fraction of the crashes happens outside the month of earnings announcements. An *et al.* (2020) recently discussed the possibility that unexpected events unrelated to strategic managerial opportunism can sometimes cause stock price crashes (*e.g.* the effect of Deepwater Horizon disaster on BP stock prices). The authors thus suggest that restricting the measurement of the crash incidence to the vicinity of earnings announcement dates can help mitigate the effect of such unexpected events.

Ergo, in model (4) we define the new dependent variable *CRASH_QEA* as taking the value of one when a stock price crash occurs withing a QEA month, and zero otherwise. Nevertheless, *ΔDR*'s coefficient remains significantly positive (p-value < 0.01). This result lends credence to the notion that the strong positive distress-crash risk relationship is probably substantiated by the bad news withholding mechanism; the kind of negative information that is held back by managers is more likely to be reflected in the earnings news on earnings announcement dates (Hutton *et al.*, 2009; Zhu, 2016; An *et al.*, 2020). As a supplementary move, in model (5) we define the new dependent *CRASH_NONQEA* as taking the value of one when a stock price crash occurs within a non-QEA month, and zero otherwise. Interestingly, *ΔDR*'s coefficient is again statistically positive (p-value < 0.01), whilst having an even higher value compared to that in model (3). While unexpected events outside the control of managers can sometimes cause stock price crashes, it is unlikely to also cause an increase in a firm's distress risk level in the three months prior to such an unexpected event. Based on this argumentation, the high positive coefficient of 0.098 for *ΔDR* in model (5), compared to the coefficient value of 0.064 in model (4), may suggests that negative information is diffused (or discovered) more intensively in the market during non-QEA months.

Next, we augment our baseline regression models by considering a broad set of additional control variables that potentially correlate with either distress risk (*ΔDR*) or the incidence of stock price crashes (*CRASH*). The list of these controls includes the following risk-, competition- and investment-related variables: the market default likelihood indicator (*MDLI*) as developed by Andreou (2015) to control for market-wide financial distress risk; the inverse current ratio (*CL/CA*) to proxy for firm-specific financial liquidity; a binary variable to control for the firm's age (*AGE_10*), set equal to one if the firm's age is lower than 10 years and zero otherwise; the ratio of goodwill to total assets (*GOODWILL*); research and development expense to total sales (*R&D/SALES*); *TOBIN'S_Q*, calculated as a firm's market value over total assets, featuring a potential proxy for mispricing; and investment-to-asset ratio (*INV*) estimated similarly to Lyandres, Sun, and Zhang (2008). We also use two competition-related variables as in Andreou *et al.* (2017b), the firm's degree of competitiveness (*COMPETITIVENESS*), calculated as the industry-adjusted ratio of firm operating profit to sales, and the firm's Herfindahl-Hirschman index (*HHI*), calculated as the squared of the firm's market share multiplied by a hundred.

These model estimations are shown in models (6) to (9). Irrespective of the model considered, the coefficient of *ΔDR* is always positive and highly statistically significant at the one percent level. Further, the firm's financial constraints (*SA*) are no longer statistically significant, casting further doubt on the notion that financial distress is merely one dimension of financial constraints. The latter also favors our view that distress risk is a contextually different crash risk determinant than financial constraints, which spans critical information regarding the firm's true state of economic fundamentals in a timelier fashion.

Overall, Table 4 provides strong evidence (both in terms of statistical and economic significance) to support our argument that short-term increases in a firm's distress risk increase the likelihood of a stock price crash incidence in the future, *i.e.* a *positive distress-crash risk relationship*.

4.2. Robustness checks

4.2.1. Alternative crash risk measures

We test the robustness of the baseline model's results by using alternative measures of stock price crashes and distress risk. In this regard, we estimate models with the five alternative crash risk measures as presented in Section 2.2, namely *CRASH_3.2*, *CRASH_20PRC*, *NCMRET*, *NCSKEW*, and *DUVOL*. We also use another alternative distress risk measure, denoted as *ΔDR_ALT*. This measure follows the study of Charitou *et al.* (2013), whereby the annualized volatility of firm assets is calculated with monthly firm value returns that are adjusted for the firm's total payout (detailed definition in the Appendix). The results are reported in Table 5. To save on space, we omit the coefficients for the control variables and tabulate only the coefficients of ΔDR_{t-1} and ΔDR_ALT_{t-1}.

First, models (4) and (7) of Table 5, Panel A provide strong evidence that a positive distress-crash risk relationship is also prevalent for the two alternative dichotomous stock price crash measures. Second, the results from all other models provide evidence to support that the alternative distress risk measure (*ΔDR_ALT*) is also strong statistically positive (p-values < 0.01) in predicting crashes across all three dichotomous measures. The models in Table 5, Panel B show a strong positive (p-value < 0.01) relationship between the distress risk measures and the continuous crash risk proxies (*NCMRET*, *NCSKEW*, and *DUVOL*). Overall, the results from this analysis suggest that our findings are robust to alternative definitions of the main dependent and independent variables.

[Insert Table 5, here]

4.2.2. Merton *DD* component variables

We then examine whether the impact of ΔDR on future stock price crashes is driven *only* by the information impounded in any of the three main components used in its calculation as per the Merton *DD* model. We re-estimate the baseline model by including in turn: (i) the stock's cumulative return, AR_{t-2};[10] (ii) the firm's market-value of assets to book-value of debt (*e.g.* the inverse of market-inferred leverage), V/D_{t-1}; and (iii) the firm's volatility of total asset returns, $\sigma_{BS,t-1}$. The results are presented in Table 6 (the baseline model's results are repeated in model (1) for comparison purposes). Intriguingly, ΔDR remains statistically positive (p-value < 0.01) after we control for these three components of distress risk.[11] These findings broadly confirm that the crash risk forecasting power of ΔDR is not driven by any of its components, but is instead the combined effect derived from its definition as per the Merton *DD* model.

[Insert Table 6, here]

4.2.3. Time span predictability of *ΔDR*

One can claim that the market data used to estimate ΔDR share common information with the stock returns associated with the estimation of stock price crashes. We measure ΔDR in month t-1 (spanning distress risk information in months from t-4 to t-1) to predict stock price crashes in month t, hence we use non-consecutive periods in our regression analyses. This choice aligns with the common practice in other crash risk studies that utilize a *lead-lagged relationship* between their dependent and explanatory variables. Nevertheless, one can suggest that our measurement points are still close to each other in time and that they might somehow be mechanically correlated. We perform certain tests to alleviate concerns that the distress risk measure encompasses market data information that might be mechanically correlated with stock price crashes.

Specifically, we re-estimate our baseline model by using changes in distress risk measured at more distant time periods in the past. We investigate the time-span predictability of distress risk and report the results in Table 7 (the baseline model's results are repeated in model (1) for comparison purposes). Model (2) examines the impact of the ΔDR as measured in month t-2 (ΔDR_{t-2}), which is the change in distress risk from month t-5 to t-2, whereas model (3) examines the impact of the ΔDR as measured in month t-3

[10] We measure AR in month t-2 because the baseline specification already includes RET_{t-1}.
[11] We also investigate model specifications that include the rolling 52-week volatility of return instead of the firm's volatility of total asset returns, to find that ΔDR_{t-1} remains positive and highly statistically significant.

(ΔDR_{t-3}), which is the change in distress risk from month *t*-6 to *t*-3. Likewise, models (4) to (6) measure ΔDR at more distant periods in the past.

Overall, the results reported in Table 7, and particularly the ones for models (2) to (4), show that the coefficients of ΔDR remain statistically positive (p-value < 0.01) up to four months ahead of the stock price crash month. This suggests that short-term changes in distress risk can predict stock price crashes as early as four months prior to the crash event, while also, in principle, providing more support for our findings in Table 4.

There is another intriguing observation from Table 7, whereby model (6) shows no statistical association between short-term changes in distress risk—when estimated more than 5 months in the past—and the following month's stock price crashes. In unreported results, we confirm the absence of a statistically positive effect when using all other ΔDR measurements up to *t*-12 (*i.e.* ΔDR_{t-7} … ΔDR_{t-12}). These latter results, in combination with the results of models (1) to (4), lend credence to our viewpoint that a firm's distress risk captures critical (short-term oriented) information regarding a firm's economic fundamentals at points in time that are close to the crash event. These results also resonate with prior studies (*e.g.* Zhu, 2016; Andreou *et al.*, 2017b) not reporting evidence in support of a positive distress-crash risk relationship; this was evidently due to the use of yearly analyses, over monthly intervals and, consequently, the inability of those studies to capture the relevant information context impounded in distress risk.

[Insert Table 7, here]

4.3. Time evolution of distress risk around the crash events

To add further nuance to our findings, we scrutinize the data to ensure that our main finding regarding the positive distress-crash risk relationship is indeed arising in the short-term and as it gets closer to the stock price crash event. This examination is based on the evolution over time of distress risk as illustrated in Figure 1. This figure presents the reaction of distress risk (*DR*) around the months that sample firms experience a stock price crash (*i.e. crash group*), as illustrated by the red line. For comparison purposes, the figure depicts the behavior of distress risk using two control-matched samples. The first control group consists of *non-crash firms* matched on the total similarity measure of Hoberg and Phillips (2016) and illustrated with the green line.[12] The second control group consists of *non-crash firms* matched by the firm's

[12] The matching is performed on a firm-by-firm basis in the fiscal year that a firm in the crash group experienced a stock price crash. The total similarity measure of Hoberg and Phillips (2016) is based on words that firms use to describe their products in 10-K annual filings, where for each firm a pairwise word similarity score is computed. The total similarity measure was collected from the Hoberg-Phillips Data Library in February of 2019 (http://hobergphillips.tuck.dartmouth.edu/industryconcen.htm). Our analysis also used the firms' size as the matching measure for each month and industry (Fama and French, 1997, 48-industry classifications), where the results emerged as quantitatively similar.

prior year's distress risk, as illustrated by the blue line.[13] The event window spans the six months preceding and then following the crash event month ($CRASH_t = 1$) observed for the firms in the crash group. Inspection of Figure 1 reveals that the mean distress risk for firms in the crash group is about 3.0 percent six months prior to the crash event, and increases enormously over time to climb close to 6.0 percent during the crash month. Two interesting findings emerge, following the crash event, the mean distress risk for these firms does not revert to its prior level; instead, it remains around 9 percent at the end of the six-month window; and the control-matched firms do not show any noticeable deviation in behavior, as there is no significant variation in distress risk during this 12-month period for these firms.

[Insert Figure 1, here]

The evidence in Figure 1 lends further credence to the existence of a positive distress-crash risk relationship, and squares with the notion that stock price crashes for firms in the crash group do not occur due to a briefly relevant piece of bad news. On the contrary, these crashes most likely happen due to the discovery of long-term negative information that was strategically withheld from the market, something that preemptively diminished the investors' perceptions about the firms' true state of economic fundamentals, causing a rather sharp positive increase in the firms' distress risk level. Presumably, such sudden increases in distress risk are associated with situations whereby managers behaved opportunistically for some time and held back bad news that had a measure of adverse effect on the firms' prospects. Further, based on our previous argumentations, restrictions on private information and various market frictions makes it unlikely that outside investors are able to accurately assess the accumulated negative information in a timely fashion, and hence to promptly adjust (the already inflated) stock prices well ahead of the crash incidence. As such, the true state of a firm's economic fundamentals, as measured by its monthly level of distress risk, is unlikely be appraised accurately during the periods that managers are incentivized and capable to hide bad news from investors. In support of this, Figure 1 shows evidence that the distress level for the crash group is below the one for the matched-control groups four or more months before the crash incidence (the same behavior is observed if this window is expanded for more than six months before the crash event).

In summary, we interpret the sharp increase in distress risk during the short period preceding the crash event as a phenomenon that is fueled by the gradual release of accumulated negative information (the so-call bad news hoarding mechanism), which is a central agency-based tenet in crash risk literature.

[13] Matching based on the distress risk 12 months before the crash is used to preclude the possibility that our findings are driven by certain characteristics related to the distress risk of crashed firms in normal periods rather than sudden increases in their distress risk in the period prior to the crash.

4.4. Endogeneity treatments

In this section, we aim to buttress our inferences regarding the strong positive distress-crash risk relationship. To do so, we employ a battery of econometric treatments to mitigate any endogeneity concerns arising from either reverse causality or unobservable heterogeneity (*i.e.* omitted variables bias).

4.4.1. Reverse causality

As a first step to fence against reverse causality issues, our regression results consistently rely on a *lead-lagged relationship*, whereas *ΔDR* measured in month *t*-1 is used to forecast stock price crashes in the subsequent month *t*. Further, we consistently include the prior year's *NCSKEW* value in the baseline control variables, to account for crash risk persistency observed in prior studies. We empirically confirm that the positive distress-crash relationship remains strong when the baseline model is re-estimated by also including the three most recent lagged values of the dependent variable ($CRASH_{t-1}$, $CRASH_{t-2}$, and $CRASH_{t-3}$). If our main inferences are confounded by reverse causality, then the inclusion of these additional lagged values would attenuate (if not eliminate) the strong positive distress-crash risk relationship. On the contrary, unreported results show that the coefficient of ΔDR_{t-1} is 0.087 (p-value < 0.01), a value that is slightly higher compared to the baseline model.[14]

Another way to test whether the positive distress-crash risk relationship is confounded by dynamic reverse causality is to swap the two main variables of interest. We conduct this empirical analysis by estimating six regression models to investigate if the most recent lagged values of stock price crashes ($CRASH_{t-1}$, $CRASH_{t-2}$, …, $CRASH_{t-6}$) are associated with changes in distress risk in the future, particularly the change in distress risk from month *t* to month *t+3* ($\Delta DR_{t,t+3}$). The results of this analysis are presented in Table 8, Panel A. Overall, the results suggest that past values of stock price crashes are not positively associated with future short-term changes in distress risk. The only exception is the positive relationship between $CRASH_{t-1}$ and $\Delta DR_{t,t+3}$ (p-value < 0.05) as shown in model (1). This is most likely an outcome of the investors' response *following* the actual crash event. For example, Chang *et al.* (2017) show that institutional investors sell a firm's shares more aggressively upon the release of bad news that causes a stock price crash. The heavy selling pressure from investors can magnify market responses to negative information about firms, suppress equity values even further, and drive firms' distress risk to higher levels. This selling pressure appears to have a short-lived effect because, as depicted in Figure 1, firms' distress risk continues to quickly increase only up to month *t*+1 and thereafter remains relatively stable. Despite the above explanation, one could interpret the results of model (1) in Table 8, Panel A as muddying our claim

[14] Our main findings remain the same if instead the baseline model is re-estimated by including either the three most recent lagged values of *NCSKEW* ($NCSKEW_{t-1}$, $NCSKEW_{t-2}$, and $NCSKEW_{t-3}$), or the three most recent lagged values of *DUVOL* ($DUVOL_{t-1}$, $DUVOL_{t-2}$, and $DUVOL_{t-3}$). These unreported results are available upon request.

regarding a causal positive distress-crash risk relationship. In response, we provide two more analyses that repudiate this concern.

[Insert Table 8, here]

First, in Table 7, we show that not only is ΔDR_{t-1} (z-stat = 8.06) strong positively related to $CRASH_t$, but the same applies when using ΔDR_{t-2} (z-stat = 5.54), or ΔDR_{t-3} (z-stat = 5.08) or even ΔDR_{t-4} (z-stat = 3.99). Hence, there is evidence to suggest that the strong positive distress-crash risk relationship is not idiosyncratic only to ΔDR_{t-1}, but substantiated with other more distant lagged values of the main explanatory variable. Conversely, the same pattern does not emerge when comparing the results from models (2) to (5) of Table 7 with corresponding models (2) to (5) in Table 8, Panel A. For reverse causality to be an explanation of the positive distress-risk relationship, one would also expect more distant lagged values of the *CRASH* variable to be associated with changes in distress risk in the future. This is not the case based on the evidence in our results.

Second, Table 8, Panel B presents additional analysis in which our baseline regression model is re-estimated after excluding all observations in the month(s) following the incidence of stock price crashes. Specifically, model (1) excludes all firm-month observations in the month that follows (t+1) a stock price crash incidence; likewise, model (2) collectively excludes all firm-month observations for the two subsequent months (t+1 and t+2), and model (6) collectively excludes all firm-month observations for the six subsequent months (t+1, t+2, …, t+6). Accordingly, if the observed positive distress-crash risk relationship is mechanically driven by distress risk increases that happen due to crash incidences, then we would expect the relationship to either attenuate or completely vanish after excluding all observations that occur after the crash month. Yet, the results across the models strongly indicate the contrary: the strong positive association between ΔDR_{t-1} and $CRASH_t$ endures and in fact appears to grow even stronger as more and more firm-month observations are excluded following the crash month.

Overall, our analyses suggest that the positive distress-crash risk relationship cannot be explained by reverse causality and cast considerable doubt that the relationship is spuriously induced.

4.4.2. Unobserved heterogeneity

Some omitted unobservable firm characteristics (fixed and time-varying) may simultaneously affect both the firm's distress risk and the occurrence of crashes; such endogeneity may confound our results and render our inferences invalid. While empirical models cannot possibly capture all the antecedents of stock price crash risk, our first effort to mitigate endogeneity issues entails controlling for fixed- and time-varying unobserved heterogeneity. In this vein, our analyses in Table 4 include time- and industry-fixed effects, along with a large array of control variables spanning a wide spectrum of relevant firm- and market-related characteristics. The message conveyed from the results of the expanded model (9) of Table 4, which includes 18 control variables, strongly supports the existence of a positive distress-crash risk relationship. Regardless, in this section, we conduct two more analyses to guard against erroneous inferences emerging from unobserved heterogeneity.

The first analysis considers the conditional firm-fixed effects logit estimator, a treatment that enables us to mitigate the concern that omitted time-invariant firm characteristics may be driving the findings. These results are shown in model (2) of Table 9. Accordingly, the inclusion of firm fixed-effects does not have any material impact on the coefficient of ΔDR_{t-1}, which maintains its statistically positive value (p-value < 0.01) and is even higher (0.093 *vs.* 0.086) compared to our baseline model (repeated in model (1) of this table for convenience). The same conclusion is reached when firm-fixed effects are included in models where we use the three continuous crash risk measures (*NCMRET*, *NCSKEW*, and *DUVOL*) as the dependent variable.[15]

[Insert Table 9, here]

Second, to further scrutinize the causal effect of short-term changes in distress risk on the following period's crash risk, we employ an instrumental variable approach. The search for a truly exogenous instrumental variable is not an easy task, and so we look to the empirical asset pricing literature. Fama and French (1992) suggest that firm size and book-to-market feature cross-sectional variation in average returns that is related to relative distress risk. Fama and French (1993) interpret the average *HML* (High-minus-Low) return as a premium for a state variable risk related to relative distress. Vassalou and Xing (2004) suggest that the *SMB* (Small-minus-Big) and *HML* returns contain potentially significant default-related information. Further, Campbell *et al.* (2008) report that financially distressed firms have high loadings on the *HML* and *SMB* factors.

In the spirit of Fama and French, the *SMB* factor is the monthly spread in returns formed by taking the difference for the portfolio return of small market capitalization *vs.* the portfolio return of big market capitalization firms. Likewise, the *HML* factor is the monthly spread in returns formed by taking the difference for the portfolio return of high book-to-market ratios *vs.* the portfolio return generated by low book-to-market ratios. To increase the granularity of the instruments, we estimate the *monthly industry return spread* for both factors by using the 48-industry classification by Fama and French (1997). Therefore, our estimations can be perceived to capture the *monthly industry–specific SMB* and *HML* information.

The two instruments, *SMB* and *HML*, reasonably satisfy the relevance condition as they are correlated with a firm's distress risk (the endogenous variable) as suggested in the literature (Fama and French, 1992, 1993; Vassalou and Xing, 2004). In our case, *SMB* and *HML* are constructed by using portfolio level

[15] The unreported corresponding coefficients and statistical significance for ΔDR_{t-1} are: 0.088 (p-value < 0.01) for *NCMRET*, 0.032 (p-value < 0.01) for *NCSKEW* and 0.023 (p-value < 0.01) for *DUVOL* (the full results are available upon request).

information emanating from industry-specific returns. Hence, both instruments should also satisfy the exclusion condition; there is no plausible argument for either *SMB* or *HML* to correlate with a firm's crash risk—a purely *idiosyncratic* firm characteristic—in any way other than the distress risk channel. Overall, the monthly, industry-defined *SMB* and *HML* factors qualify as valid instruments to enable us to mitigate endogeneity concerns in distress risk.

To implement the instrumental approach, we follow two stages: in the first stage, we separately regress firms' distress risk on the monthly, industry-defined *SMB* and *HML* factors. In the second stage, we compute the instrumented short-term change in distress risk by using the fitted values from the first-stage regression. The results for the first-stage regressions (not reported for brevity) show that both instruments are significantly related (p-values < 0.01) to distress risk, whereas the resulting adjusted R^2 ($\cong 0.28$) and F-statistic (p-value $\cong 0$) suggest that the model does not suffer from the issue of weak instruments.

Model (3) of Table 9 shows the second-stage results when using the monthly, industry-defined *SMB* factor as the instrument ($\Delta DR\_IV\_SMB_{t-1}$), model (4) shows the results when using the month-industry defined *HML* factor as the instrument ($\Delta DR\_IV\_HML_{t-1}$), and model (5) shows when both *SMB* and *HML* are used as instruments ($\Delta DR\_IV\_SMB\&HML_{t-1}$). All instrumental analysis results are consistent with our baseline model inferences, supporting a causal positive distress-crash risk relationship.

4.4.3. The Sarbanes-Oxley Act as a quasi-experimental setting

To further address endogeneity concerns, we examine the extent to which the Sarbanes-Oxley Act of 2002 (SOX)—an exogenous regulatory event that occurred in the course of our sample period—may have influenced managers' practices with respect to their tendency to release bad news through formal corporate disclosure. For instance, Cohen, Dey, Lys, and Sunder (2007) document evidence consistent with firms having less flexibility in reporting earnings-increasing discretionary accruals in the post-SOX period. The latter is further corroborated by Hutton *et al.* (2009), who report that accounting opacity associated with earnings management (*i.e.* managerial effort to hoard negative information) has declined in the post-SOX years, whilst Callen and Fang (2017) discuss that the enactment of SOX has attenuated withholding of bad news and improved managerial disclosure and transparency.

In the post-SOX period, concerns around litigation increased, with managers subject to stricter monitoring from auditors, creditors, and other stakeholders. This should have also limited self-interested behavior like engaging in income-increasing practices through persistently withholding bad news. It is thus reasonable to assume that the enactment of SOX, due to its associated negative externalities (*i.e.* increased litigation risk, loss of reputation, legal actions, *et cetera*), has decreased managers' willingness to repeatedly withhold bad news from being included in their firms' quarterly earnings statements. Ergo, SOX can be used as a quasi-experimental setting enabling us to investigate this differential effect on the incremental information content of quarterly earnings announcements and, *subsequently*, its implications for the distress-crash risk relationship.

Ceteris paribus, relative to the pre-SOX period, in the post-SOX period we expect to observe a heightened incidence of stock price crashes occurring in the quarterly earnings announcement (QEA) months. To investigate this proposition, we estimate the following model:

$$CRASH_{i,t} = \delta_0 + \delta_1 DPOST\_SOX + \delta_2 DQEA_t + \delta_3 DPOST\_SOX \times DQEA_t + \sum_{k=1}^{K} a_k CONTROLS_{i,t-1} + e_{i,t}, \quad (10)$$

where *DPOST_SOX* is a binary variable that takes the value of one in the period from July 2002 to June 2005 (POST-SOX), and zero in the period from July 1999 to June 2002 (PRE-SOX), *DQEA* takes the value of one for months that a firm is making a quarterly earnings announcement, and zero otherwise, while *CONTROLS* is the array of covariates used in our baseline model as per *Eq.* (9). We estimate the above model by respectively considering the three-year period before and after the SOX enactment in July 2002 to make sure that we include periods that feature important and relevant information associated with the event. In practical terms, *Eq.* (10) is like a difference-in-differences regression model, whereby the earnings announcement months (captured with *DQEA*) feature the treated observations, and the interaction term *DPOST_SOX×DQEA* captures the incremental difference in stock price crashes between the *QEA* months and control observations (*i.e.* non-QEA months) after the enactment of the SOX.

[Insert Table 10, here]

We report the results in Panel A of Table 10. As shown, the coefficient on the interaction term *DPOST_SOX×DQEA* is statistically positive (p-value < 0.01), suggesting that, relative to the non-QEA months where no firm disclosure is released, the QEA months in which managers release the quarterly financial statements exhibit a notably higher incidence of stock price crashes in the POST-SOX period.

The above results lend credence to the notion that the POST-SOX-QEA months are associated with an increased disclosure of bad news that cause more frequent stock price crashes. This provides us an identification setting to vindicate the positive distress-crash risk relationship. For short-term changes in distress risk to have a positive causal effect on the following period's crash risk, we should also observe that ΔDR_{t-1} in the POST-SOX-QEA months are notably higher compared to the PRE-SOX-QEA months. These results are presented in Panel B of Table 10 and support the expectations.

Specifically, conditioning on a stock price crash happening in month t ($CRASH_t = 1$), the mean value of ΔDR_{t-1} for the PRE-SOX-QEA months is -0.58 percent and increases to 0.75 percent for the POST-SOX-QEA months, with the difference of 1.32 percent being statistically significant (t-statistic = 2.15) [column (3) upper part of Panel B]. Evidently, the enactment of SOX has resulted in an increasing shock in distress risk in the QEA months. Looking at the counterpart control observations comprised by the non-QEA months, we do not observe the same behavior: conditioning on a stock price crash happening in month t

($CRASH_t$ = 1), the difference in the mean values of ΔDR_{t-1} in the POST-SOX *vs.* PRE-SOX periods is only 0.70 percent and indistinguishable from zero (*t*-statistic = 1.01) [column (3) bottom part of Panel B]. Based on our prior argumentation, we would not expect the SOX enactment to have any impact on managers' tendency to withhold (or release) bad news during the non-QEA months. Overall, these comparisons suggest that ΔDR_{t-1} increases considerably in the POST-SOX-QEA, whereby firms experienced an increasing incidence of stock price crashes triggered by the increasing quantity of bad news released by managers who were compelled to do so under the more stringent SOX environment.

5. Additional analyses

5.1. Distress risk, earnings management practices, and the bad news hoarding mechanism

Identifying the underlying reason for which we observe a positive distress-crash risk relationship is important to enable us to substantiate that short-term changes in distress risk span critical information pertaining to the withholding of bad news initiated by managers to mask their firms' true state of fundamentals. Following the growing literature in this area, we investigate whether the positive distress-crash risk relationship is mediated through management's manipulation practices aimed at inflating earnings. The crash risk literature demonstrates that managers primarily rely on earnings manipulation to facilitate bad news hoarding behavior that services their career and wealth-related concerns (*e.g.* Hutton *et al.*, 2009; Kothari *et al.*, 2009; Kim *et al.*, 2011a; Callen and Fang, 2015; Kim and Zhang, 2016; Zhu, 2016; Andreou *et al.* 2017b; Chen *et al.* 2017).

Following the literature, we assume that aggressive earnings management is likely to proxy for management's general proclivity to hide information from the capital market to retain high market valuations. We thus estimate a measure of accounting opacity capturing financial reporting quality based on an indicator of earnings management, whereby information opacity is viewed as an outcome of managerial opportunism that is exercised through the manipulation of a firm's discretionary accruals.

We carry the investigation in two steps. First, we investigate whether short-term changes in distress risk are associated with future short-term changes in opacity. Specifically, we estimate the following recursive regression:

$$\Delta OPACITY\_3M_{i,t}=\theta_0+\theta_{1,t}\Delta DR_{i,t-1}+\sum_{k=2}^{K-1}\theta_{k,t}CONTROLS\_NO\_OPACITY_{i,t-1}+\varepsilon_{i,t} \quad (11)$$

where $\Delta OPACITY\_3M_{i,t}$ measures the change in the accounting opacity variable from month *t-3* to *t*, and the baseline group of controls is as per *Eq*. (9), excluding the accounting opacity variable. In the second step described below, we investigate if the short-term changes in distress risk embed important earnings management information. For this reason, following the empirical approach in prior asset pricing studies (Welch and Goyal, 2008), *Eq*. (11) is estimated recursively in an out-of-sample fashion that avoids potential overfitting problems, and uses data that would only be accessible to an investor when making a real-time assessment of a situation (we use 12 months of data as the initial estimation period).

The first-step results are shown in Panel A of Table 11, whereby we observe that ΔDR_{t-1} is positively associated with $\Delta OPACITY\_3M_t$ (p-values < 0.01). Overall, these results show that short-term changes in distress risk associate with critical information relating to situations whereby managers attempt to camouflage bad news through earnings management manipulations. This evidence is very important because it qualifies the short-term changes in distress risk as a predictor of the (short-term and incremental) efforts of managers to disguise their firms' true state of fundamentals through income-increasing practices.

Second, to complete our investigation of whether the positive relationship between ΔDR_{t-1} and $CRASH_t$ is mediated by the management's earnings management practices, Panel B of Table 11 investigates the relationship between the portion of short-term changes in distress risk attributed to accounting opacity ($\Delta DR\_\widehat{OPACITY}\_3M_{t-1}$) and the following month's stock price crashes. Since $\Delta DR\_\widehat{OPACITY}\_3M_{t-1}$ is a predicted quantity computed using the recursively estimated coefficients from *Eq*. (11), it underpins the ability of ΔDR_{t-1} to detect earnings manipulations aiming to camouflage a firm's (true state of) economic fundamentals. The results show that $\Delta DR\_\widehat{OPACITY}\_3M_{t-1}$ is positively associated with $CRASH_t$, supporting the notion that management's bad news hoarding behavior is a potential channel through which financial distress risk affects future stock price crashes.

[Insert Table 11, here]

5.2. The moderating role of information asymmetry

Stock price crashes are likely to occur among firms facing high agency problems. Such agency problems arise because self-interested managers tend to exploit information asymmetries that exist between managers and shareholders by concealing negative information and engaging in short-sighted price maximization that better serves their own interests (Kothari *et al.*, 2009; Hutton *et al.*, 2009; Kim *et al.*, 2011b). In this vein, there is plenty of evidence to support that a firm's information environment plays a key moderating role in the context of crash risk, because firms with high information asymmetry are likely to suffer more from severe agency conflicts between insiders and outsiders than those with low information asymmetry (Callen and Fang, 2015; Andreou *et al.*, 2017a). We therefore examine whether information asymmetry moderates the distress-crash risk relationship.

For this purpose, we rely on information related to financial analysts, primarily because they embody an external monitoring mechanism capable of reducing information asymmetries between managers and shareholders (Lang, Lins, and Miller, 2003). Because of their additional role in information intermediation, analysts not only monitor managers directly, they can also reduce information asymmetry between managers and investors, which, in turn, facilitates external monitoring by outside investors (Chen *et al.*, 2017; Kim *et al.*, 2019). In general, financial analysts contribute towards closing the information asymmetry gap between corporate insiders and outside equity investors, something that helps to curb managerial opportunism and mitigate bad incentives for managers to accumulate negative information. Within the agency context of our investigation, stock price crashes should be more prevalent in firms facing high information asymmetry environments in which managers have greater opportunities to manipulate outside investors' expectations on the firm's economic fundamentals (by strategically withholding bad news relating to adverse performance outcomes).

Accordingly, we use two measures of information asymmetry derived from financial analyst information, specifically, analyst coverage and analyst earnings forecasts dispersion (*AFD*) computed each month. Analyst coverage is the number of analysts who follow a particular firm each month (firms with missing analyst data are recorded as zero coverage). *AFD* is defined as the standard deviation of earnings-per-share forecasts for the current fiscal year over the absolute value of the mean earnings forecasts (*e.g.* Diether, Malloy, and Scherbina, 2002; Callen and Fang, 2015). For this measure, we require at least two analysts to follow the firm and zero standard deviation values are excluded.

Table 12 presents the logit regression estimates based on five subsamples derived from these two information asymmetry measures. Specifically, models (1) to (3) show the results for firm-month subsamples formed when firms are sorted in terciles based on analyst coverage in month *t*, whereby we assume that model (1) includes monthly observations featuring the highest level of information asymmetry, and model (3) includes monthly observations featuring the lowest level of information asymmetry in our sample (and model (2) is in between). In further support of the hoarding of bad news mechanism, we would expect the impact of ΔDR_{t-1} to be greatest in model (1), smallest in model (3) and its impact to be in-between for model (2). As shown in Table 12, the impact of ΔDR_{t-1} is more pronounced in model (1) when there is a high level of information asymmetry while the impact of ΔDR_{t-1} disappears when the information asymmetry is at its lowest level in model (3).

[Insert Table 12, here]

Furthermore, models (4) and (5) provide another partition of our sample, whereby information asymmetry is measured using the analysts' forecast dispersion, with model (4) including observations featuring low levels (*i.e.* below median *AFD*) and model (5) including observations featuring high levels (*i.e.* above median *AFD*) of information asymmetry. The empirical results suggest that the impact of ΔDR_{t-1} on future stock price crashes is present only in model (5).

Overall, the evidence in Table 12 lends further support that the positive distress-crash risk relationship is substantiated by the bad news hoarding mechanism fueled by high information asymmetry environments.

5.3. Further evidence that short-term changes in distress risk capture bad news hoarding efforts

In the analyses above, we find that the short-term changes in distress risk span critical information relating to management's efforts to conceal bad news. Our inferences suggest that ΔDR_{t-1} is a successful market-based predictor of stock price crashes because it detects, in a timely manner, situations in which managers are disguising and hoarding bad news.

We provide supplementary analysis to further support this working hypothesis that short-term changes in distress risk timely capture the negative information that gradually spills into the market ahead of the crash event. Specifically, we investigate the relation between ΔDR_{t-1} and $CRASH_t$ under the lens of information relating to financial analysts' buy-sell revisions. These revisions can hit the market at any point in time, not just on the quarterly earnings announcement dates. Analysts can proceed with an upward (*i.e.* positive) or downward (*i.e.* negative) revision of their recommendations when they recognize some key fundamental information that will affect the company's market value in the near future.

Extant literature generally suggests that analysts possess high-level financial skills and information-searching ability, and, hence, their outputs provide value to capital market participants, principally through their information discovery and intermediary role (*e.g.* Huang, Lehavy, Zang, and Zheng, 2018; Kim, Lu, and Yu, 2019).[16] Primarily through processing, uncovering, and disseminating new information, financial analysts bring incremental information to the market, empowering investors to preemptively decipher managerial opportunism situations and identify whether managers are hoarding bad news. Hence, analysts' revisions offer an opportunity to investigate whether the short-term changes in distress risk promptly impound critical information coming from a different, albeit important, information channel.

It is natural to assume that negative revisions convey a worsening of the analysts' expectations regarding the firm's state of economic fundamentals. In principle, these revisions can be driven by a variety of newly discovered negative information that downgrades the firm's prospects (*e.g.* entry of new competitor, drop in sales, *et cetera*). In such situations, a negative revision may trigger a sell-off that depresses the stock price, but is unlikely to systematically trigger a stock price crash. This is not however the case when the negative revisions are triggered by analysts discovering that crucial negative information had been strategically concealed by managers from the market with the sole purpose of presenting more favorable economic fundamentals to investors. In this case, the analysts' negative revisions will most likely be

[16] According to Huang *et al.* (2018), the analyst information discovery role includes personal research efforts to collect and generate information that is otherwise not readily available to investors. Such efforts aim to generate *new information signals* regarding a firm's prospects, and involve, among others, personal research and channel checks, private interactions with a firm's top management team at the headquarters and division units, processing of information collected from various sources (*e.g.* other information intermediaries, peer firms in the industry, and independent research agencies), and undertaking original analysis by "connecting the dots". Kim *et al.* (2019) discuss that their information intermediary role is relevant to investors' assessment of a firm's downside risk, mainly because analysts engage in activities that facilitate the propagation of bad news in the market.

accompanied with a stock price crash event. This distinction allows us to conduct another analysis to test the working hypothesis, whereby increases of ΔDR_{t-1} convey information for imminent crash risk problems fueled by managerial opportunism. Therefore, a situation where negative revisions trigger a stock price crash represents an ideal setting to explicitly investigate whether bad news hoarding drives the relationship between ΔDR_{t-1} and $CRASH_t$.

In terms of the empirical analysis that is shown in Table 13, a negative revision is when, between two consecutive months, there is either an increase of the average "sell" percentage recommendation (Panel A) or a decrease in the average "buy" percentage recommendation (Panel B). To examine the impact on ΔDR_{t-1} on following month's stock price crashes conditional on information relating to the negative revision, we use three different subsamples: (i) all monthly observations for which we observe a negative revision (upper part of Table 13), (ii) monthly observations for which we observe a negative revision that falls *within* the quarterly earnings announcement (QEA) months (middle part of Table 13), and (iii) monthly observations for which we observe a negative revision that falls *outside* the quarterly earnings announcement (non-QEA) months (bottom part of Table 13). For each sample, we split observations featuring a negative revision in month *t* in those that co-occurred with a stock price crash event [column (1)] and those that fall in months with no crash event [column (2)], and compare the associated average value of ΔDR_{t-1} [column (3)].

As such, we expect ΔDR_{t-1} to be higher where analysts' negative revisions have caused a stock price crash (as opposed to when they have not), with the former situation indicating that the trigger for the negative revision was adverse information associated with bad news. In general, this is the pattern we observe in Table 13. For instance, regarding the whole sample analysis (upper part), there are 1,015 monthly observations for which we observe that a negative sell-related recommendation in month *t* co-occurs with a crash event [column (1)], whilst 25,760 monthly observations with a negative sell-related recommendation in month *t* are associated with no crash events [column (2)]. The corresponding monthly average values of ΔDR_{t-1} is 2.56 percent for revisions that cause a crash and 0.55 percent for revisions that do not. The difference of these two values is 2.01 percent [column (3)], which is highly statistically significant (p-value < 0.01) [column (4)]. Clearly, the cases where a negative revision is observed with no crash event outnumbers the cases that negative revisions co-occur with a crash. To remedy this, we also define a control-matched sample of *non-crash firms* based on the nearest market capitalization with *crash firms* in each industry. In the case under discussion, we are able to find 369 matching firm-month observations, whereby the monthly average value of ΔDR_{t-1} for revisions that do not cause a crash is again only 0.53 percent (with the difference to be 2.03 percent and highly statistically significant). Also, evidence in Panel A is corroborated with evidence in Panel B showing similar analysis that is conditional on negative revisions following a decrease in the analysts' "buy" recommendations.

It is noteworthy that the differences [column (3)] are higher during the non-QEA months (bottom part), something that indicates that negative analyst revisions are more important during periods that the market does not have any direct information coming from management, primarily in the form of new quarterly financial statements. This evidence squares with prior conjectures, as for example in Huang *et al.* (2018), that analysts increase their discovery efforts in periods that do not coincide with any firm disclosures, when managers have more capacity to withhold information.

Collectively, the comparisons suggest that short-term changes in distress risk in month *t*-1, captured by ΔDR_{t-1}, are particularly heightened within the group of firms that experienced a stock price crash in month *t* that analysts released a negative recommendation. These cases most likely correspond to situations where managers have strategically concealed bad news from investors for long periods. Hence, the evidence in Table 13 lends credence to the notion that increases of ΔDR_{t-1} span critical information relating to the as yet undisclosed bad events that managers are strategically concealing from the market.

[Insert Table 13, here]

6. Conclusion

Studies have so far failed to document a significant positive relationship between distress risk and future crash risk. This could be attributed to the measurement of distress risk using yearly intervals, which is the mainstream approach in the literature. We circumvent the above limitation by investigating the relationship between distress risk and future stock price crashes using monthly observation intervals. Equipped with large-scale panel data comprising 462,678 monthly observations of US-listed firms, we show that a short-term increase in the firm's distress risk leads to a higher probability of stock price crashes in the following month. Our findings are robust after controlling for 18 relevant covariates and the use of an alternative measure of financial distress and five alternative measures of crash risk. They are also robust to a range of tests to buttress against endogeneity concerns, including reverse causation treatments, instrumental variable estimations, and the use of Sarbanes–Oxley Act of 2002 as a quasi-experimental setting. Furthermore, the main findings have practical value: short-term changes in distress risk exhibit an economic significance impact that is much bigger than the impact of other prominent crash risk determinants, such as financial reporting opacity, trading volume turnover proxying for investor heterogeneity, *et cetera*.

Our findings are consistent with agency-related theoretical underpinnings, according to which a stock price crash is driven predominantly by practices used by managers to hoard unfavorable news for long periods. In line with this argument, we provide convincing evidence that withholding bad news is the underlying reason for the positive link between short-term changes in distress risk and future stock price crashes. These findings empirically confirm that distress risk conveys critical information about a firm's fundamentals pertinent to imminent stock price crash problems.

Our study has practical implications related to the fact that crash risk cannot be diversified by shareholders. Asset pricing researchers are starting to recognize the importance of crash risk (*i.e.* information for the third moment) as a determinant of stock (or option) returns, in addition to stock volatility (Yan, 2011; Conrad *et al.*, 2013; Jang and Kang, 2019). Unlike volatility risk, which can be reduced via portfolio diversification, crash risk cannot be diversified away (Guiseo and Jappelli, 2008; Abreu and Mendes, 2010; Barber and Odean, 2013). For small (usually under-diversified) investors in particular, idiosyncratic stock price crashes would translate to significant reductions in the value of their portfolios, especially if they maintain short-term investment horizons. Institutional investors with active fund managers, on the other hand, can improve their performance by monitoring their stocks for sudden short-term increases in distress risk, thus detecting bad news hoarding activities by firms whose stock they can sell in anticipation of a stock price crash. Our findings suggest that investors should assess a firm's monthly distress risk to determine—*ex ante*—its propensity towards stock price crashes.

Our results also caution investors about the association between crash risk and earnings smoothing. In the real world, managers have plenty of channels at their disposal (*e.g.* accrual manipulation, off-balance sheet items, vague company announcements, *et cetera*) to hide bad news around adverse firm performance. Extant literature admits that realized stock price crashes represent a comprehensive market-based measure spanning *all kinds of managerial efforts* to intentionally obfuscate poor performance by concealing negative information over long periods (Kim and Zhang, 2016; Callen and Fang, 2017; Andreou *et al.*, 2017b). In this regard, by relating distress risk to future stock price crashes, we provide further insights regarding the role of distress risk in detecting situations where managers exploit the bad news hoarding mechanism for their own benefit.

Appendix

Variable Definitions

This table provides detailed variable definitions. Because the analysis is implemented in a monthly frequency, the accounting data that are used are annualize-adjusted when necessary. Specifically, the accounting data from income and cash flows statements are annualized by taking the summation of the four most recent quarterly results. This process is also known as a trailing twelve-month treatment, which allows to exploit all the available information from each accounting variable for each month. Balance sheet items do not need any adjustment. To preclude look-ahead bias, market-based data are matched with quarterly accounting data by lagging them by three months.

Variable	Definition
Panel A: Dependent variables	
CRASH	An indicator variable set equal to one for month t if within this month the firm's stock experiences at least one firm-specific weekly return that falls more than 3.09 standard deviations below the mean firm-specific weekly return over the estimation period. The firm-specific weekly returns are estimated based on *Eq.* (2) following the input from the expanded index model as per *Eq.* (1) using a 52-week rolling estimation window.
CRASH_3.2	An indicator variable set equal to one for month t if within this month the firm's stock experiences at least one firm-specific weekly return that falls more than 3.20 standard deviations below the mean firm-specific weekly return over the estimation period. The firm-specific weekly returns are estimated based on *Eq.* (2) following the input from the expanded index model as per *Eq.* (1) using a 52-week rolling estimation window.
CRASH_20PRC	An indicator variable set equal to one for month t if within this month the firm's stock experiences at least one crash week. The crash week is defined as an extreme negative market-adjusted weekly return lower than -20%. The market weekly return is computed using the CRSP value-weighted return (CRSP item "vwretd").
NCMRET	The negative of the minimum market-adjusted weekly return over the most recent 26 weeks divided to the standard deviation of the market-adjusted weekly returns for the previous period. The firm's market-adjusted weekly return is defined as the difference between the stock and CRSP value-weighted weekly return.
NCSKEW	The negative conditional skewness estimated as the negative of the third moment of firm-specific weekly returns divided by the standard deviation of weekly returns raised to the third power. The firm-specific weekly returns are estimated based on *Eq.* (2) following the input from the expanded index model as per *Eq.* (1) using a 52-week rolling estimation window.
DUVOL	The "down-to-up" volatility that is equal to the 52-week rolling log difference volatilities between the negative and positive firm-specific weekly returns. The firm-specific weekly returns are estimated based on *Eq.* (2) following the input from the expanded index model as per *Eq.* (1) using a 52-week rolling estimation window.

Panel B: Merton distance-to-default (*DD*) related information	
D	The face value of debt that is estimated using the debt in current liabilities (Compustat item "dlcq") plus half the long-term debt (Compustat item "dlttq"). If the firm's D is not available, the estimation follows Campbell *et al*. (2008): D = median(D/TL)×TL, where TL stands for total liabilities, and the median is measured for the whole dataset. If D = 0, we use D = median(D/TL)×TL, where the median is now calculated only for small but non-zero values of D ($0 < D < 0.01$).
ME	The market value of equity that is equal to the stock price (CRSP item "prc") multiplied by the number of shares outstanding (CRSP item "shrout").
V	The firm's total asset value equals the firm's market value of equity (ME) plus the face value of debt (D).
AR	Annualized stock return derived from rolling 12-month cumulative returns.
σ_{BS}	The firm's volatility of total assets returns used in the estimation of DD as per *Eq*. (6). It is estimated as follows: $$\sigma_{BS}=\left(\frac{ME}{ME+D}\right)\sigma_E+\left(\frac{D}{ME+D}\right)\sigma_D,$$ where σ_E is the annualized equity volatility derived from monthly equity returns adjusted for cash dividends over a 36-month window, while σ_D is the debt volatility estimated using an approximation formula: $\sigma_D=0.05+0.25\sigma_E$.
DD	The "naïve" model's distance to default value of Bharath and Shumway (2008) as given by *Eq*. (6).
DR	The probability to default based on the Merton DD model as per *Eq*. (7).
ΔDR	The 3-month change in distress risk (DR) as per *Eq*. (8).
Panel C: Main control variables	
SIZE	The natural logarithm of the firm's market capitalization (ME).
M/B	The ratio of the firm's market capitalization (ME) over the book value of common equity (Compustat item "ceqq")
LEV	The ratio of total liabilities (Compustat item "ltq") to total assets (Compustat item "atq").
ROA	The ratio of net income (Compustat item "niq") to total assets.
SA	The firms' financial constraints proxied by the Size-Age index as developed by Hadlock and Pierce (2010).
RET	The firm's 3-month cumulative returns.
DTURN	The detrended turnover that is equal to the mean monthly turnover of the previous 6 months, detrended by the mean of turnover in the prior 18 months.
OPACITY	Financial reporting opacity estimated similar to Hutton *et al.* (2009) as the 36-month moving sum of the absolute discretionary accruals. Discretionary accruals are estimated based on the modified model of Jones (1991) using quarterly accounting variables, whereby the quarterly estimated values are used for all months spanning a fiscal quarter.

Panel D: Other variables	
MDLI	Market default likelihood index that is equal to the aggregate firm-specific probability to default as in Andreou (2015). *MDLI* is calculated as the mean value of probability to default for all non-financial firms included in the S&P 500 index. The probability to default for each firm is estimated using the Merton (*DD*) model.
CL/CA	The firm's inverse current ratio, which is equal to current liabilities (Compustat item "lctq") to current assets (Compustat item "actq").
AGE_10	A binary variable set equal to one if the firm's age is smaller than 10 years (since the firm's listing in Compustat) and zero otherwise.
GOODWILL	The ratio of goodwill (Compustat item "gdwlq") to total assets. Missing observations are replaced with zero.
HHI	The firm's Herfindahl-Hirschman index is defined as the squared of the firm's market share, *i.e.* firm's sales divided by industry's (2-digit SIC) total sales (multiplied by a hundred).
COMPETITIVENESS	The industry-adjusted ratio of firm operating profit (Compustat item "oibdp") to sales (Compustat item "saleq").
R&D/SALES	The ratio of research and development expense (Compustat item "xrdq") to total assets. Missing observations are replaced with zero.
TOBIN'S_Q	The ratio of the company's market value (*ME + total liabilities*) divided by the firm's total assets.
INV	Investment to assets ratio is estimated as the change in gross property, plant, and equipment (Compustat item "ppentq"), plus the change in inventories (Compustat item "invtq"), scaled by total assets
SMB	The Small-minus-Big factor is the monthly spread in returns formed by taking the difference for the portfolio return of small market capitalization *vs.* the portfolio return of big market capitalization firms within each industry using the 48-industry classification by Fama and French (1997).
HML	The High-minus-Low factor is the monthly spread in returns formed by taking the difference for the portfolio return of high book-to-market ratios (value stocks) vs. the portfolio return generated by low book-to-market ratios (growth stocks) within each industry using the 48-industry classification by Fama and French (1997).
AFD	Analysts' forecasts dispersion is defined as the standard deviation of all outstanding earnings-per-share forecasts for the current fiscal year over the absolute value of the mean outstanding earnings forecasts.
DPOST_SOX	An indicator variable that takes the value of one in the period from July 2002 to June 2005, and zero in the period from July 1999 to June 2002, with July 2002 to represent the month when the Sarbanes-Oxley (SOX) Act came into force.
DQEA	An indicator variable that equals one if in a particular month a firm announces its quarterly results, and zero otherwise. To define this variable, we use the quarterly reported date provided by Compustat (item "rdq").

References

Abreu, M., & Mendes, V. (2010). Financial literacy and portfolio diversification. *Quantitative Finance*, *10*(5), 515–528.

Ak, B. K., Rossi, S., Sloan, R., & Tracy, S. (2016). Navigating Stock Price Crashes. *The Journal of Portfolio Management*, *42*(4), 28–37.

An, Z., Chen, C., Naiker, V., & Wang, J. (2020). Does media coverage deter firms from withholding bad news? Evidence from stock price crash risk. *Journal of Corporate Finance*, *64*, 101664.

Andreou, C. K., Lambertides, N., & Panayides, P. M. (2021). Distress Risk Anomaly and Misvaluation. *The British Accounting Review*, 100972.

Andreou, P. C. (2015). Effects of market default risk on index option risk-neutral moments. *Quantitative Finance*, *15*(12), 2021–2040.

Andreou, P. C., Cooper, I., Louca, C., & Philip, D. (2017). Bank loan loss accounting treatments, credit cycles and crash risk. *British Accounting Review*, *49*(5), 474–492.

Andreou, P. C., Louca, C., & Petrou, A. P. (2017). CEO age and stock price crash risk. *Review of Finance*, *21*(3), 1287–1325.

Barber, B. M., & Odean, T. (2013). The Behavior of Individual Investors. In *Handbook of the Economics of Finance* (Vol. 2, pp. 1533–1570). Elsevier.

Benmelech, E., Kandel, E., & Veronesi, P. (2010). Stock-based compensation and ceo (dis)incentives. *Quarterly Journal of Economics*, *125*(4), 1769–1820.

Bharath, S. T., & Shumway, T. (2008). Forecasting default with the Merton distance to default model. *Review of Financial Studies*, *21*(3), 1339–1369.

Bleck, A., & Liu, X. (2007). Market transparency and the accounting regime. *Journal of Accounting Research*, *45*(2), 229–256.

Callen, J. L., & Fang, X. (2013). Institutional investor stability and crash risk: Monitoring versus short-termism? *Journal of Banking and Finance*, *37*(8), 3047–3063.

Callen, J. L., & Fang, X. (2015a). Religion and Stock Price Crash Risk. *Journal of Financial and Quantitative Analysis*, *50*(1), 169–195.

Callen, J. L., & Fang, X. (2015b). Short interest and stock price crash risk. *Journal of Banking and Finance*, *60*, 181–194.

Callen, J. L., & Fang, X. (2017). Crash Risk and the Auditor–Client Relationship. *Contemporary Accounting Research*, *34*(3), 1715–1750.

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

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

Charitou, A., Dionysiou, D., Lambertides, N., & Trigeorgis, L. (2013). Alternative bankruptcy prediction models using option-pricing theory. *Journal of Banking and Finance*, *37*(7), 2329–2341.

Charitou, A., Lambertides, N., & Trigeorgis, L. (2007). Earnings behaviour of financially distressed firms: The role of institutional ownership. *Abacus*, *43*(3), 271–296.

Chava, S., & Jarrow, R. A. (2004). Bankruptcy Prediction with Industry Effects. *Review of Finance*, *8*, 537–569.

Chen, J., Hong, H., & Stein, J. C. (2001). Forecasting crashes: Trading volume, past returns, and conditional skewness in stock prices. *Journal of Financial Economics*, *61*(3), 345–381.

Cohen, D. A., Dey, A., Lys, T. Z., & Sunder, S. V. (2007). Earnings announcement premia and the limits to arbitrage. *Journal of Accounting and Economics*, *43*(2–3), 153–180.

Conrad, J., Dittmar, R. F., & Ghysels, E. (2013). Ex Ante Skewness and Expected Stock Returns. *Journal of Finance*, *68*(1), 85–124.

DeAngelo, H., DeAngelo, L., & Skinner, D. J. (1994). Accounting choice in troubled companies. *Journal of Accounting and Economics*, *17*(1–2), 113–143.

Deng, X., Gao, L., & Kim, J. B. (2020). Short-sale constraints and stock price crash risk: Causal evidence from a natural experiment. *Journal of Corporate Finance*, *60*, 101498.

Diether, K. B., Malloy, C. J., & Scherbina, A. (2002). Differences of opinion and the cross section of stock returns. *The Journal of Finance*, *57*(5), 2113–2141.

Driscoll, J. C., & Kraay, A. C. (1998). Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data. *Review of Economics and Statistics*, *80*(4), 549–560.

Fama, E. F., & French, K. R. (1992). The Cross-Section of Expected Stock Returns. *The Journal of Finance*, *47*(2), 427–465.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*(1), 3–56.

Fama, E. F., & French, K. R. (1997). Industry costs of equity. *Journal of Financial Economics*, *43*(2), 153–193.

French, K. R., Schwert, G. W., & Stambaugh, R. F. (1987). Expected stock returns and volatility. *Journal of Financial Economics*, *19*(1), 3–29.

Guiseo, L., & Jappelli, T. (2008). *Financial Literacy and Portfolio Diversification*. European University Institute. Retrieved from https://cadmus.eui.eu/handle/1814/9811

Hadlock, C. J., & Pierce, J. R. (2010). New evidence on measuring financial constraints: Moving beyond the KZ index. *Review of Financial Studies*, *23*(5), 1909–1940.

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

Hoberg, G., & Phillips, G. (2016). Text-Based Network Industries and Endogenous Product Differentiation. *Journal of Political Economy*, *124*(5), 1423–1465.

Hong, H., & Stein, J. C. (1999). A unified theory of underreaction, momentum trading, and overreaction in asset markets. *Journal of Finance*, *54*(6), 2143–2184.

Hong, H., & Stein, J. C. (2003). Differences of Opinion, Short-Sales Constraints, and Market Crashes. *Review of Financial Studies*, *16*(2), 487–525.

Huang, A. H., Lehavy, R., Zang, A. Y., & Zheng, R. (2018). Analyst information discovery and interpretation roles: A topic modeling approach. In *Management Science* (Vol. 64, pp. 2833–2855). INFORMS Inst.for Operations Res.and the Management Sciences.

Hutton, A. P., Marcus, A. J., & Tehranian, H. (2009). Opaque financial reports, R2, and crash risk. *Journal of Financial Economics*, *94*(1), 67–86.

Jang, J., & Kang, J. (2019). Probability of price crashes, rational speculative bubbles, and the cross-section of stock returns. *Journal of Financial Economics*, *132*(1), 222–247.

Jin, L., & Myers, S. C. (2006). R2 around the world: New theory and new tests. *Journal of Financial Economics*, *79*(2), 257–292.

Jones, J. J. (1991). Earnings Management During Import Relief Investigations. *Journal of Accounting Research*, *29*(2), 193.

Kim, J. B., Li, Y., & Zhang, L. (2011a). CFOs versus CEOs: Equity incentives and crashes. *Journal of Financial Economics*, *101*(3), 713–730.

Kim, J. B., Li, Y., & Zhang, L. (2011b). Corporate tax avoidance and stock price crash risk: Firm-level analysis. *Journal of Financial Economics*, *100*(3), 639–662.

Kim, J. B., Lu, L. Y., & Yu, Y. (2019). Analyst coverage and expected crash risk: Evidence from exogenous changes in analyst coverage. *Accounting Review*, *94*(4), 345–364.

Kothari, S. P., Shu, S., & Wysocki, P. D. (2009). Do managers withhold bad news. *Journal of Accounting Research*, *47*(1), 241–276.

Lang, M. H., Lins, K. V., & Miller, D. P. (2003). ADRs, analysts, and accuracy: Does cross listing in the United States improve a firm's information environment and increase market value? *Journal of Accounting Research*, *41*(2), 317–345.

Lyandres, E., Sun, L., & Zhang, L. (2008). The new issues puzzle: Testing the investment-based explanation. *Review of Financial Studies*, *21*(6), 2825–2855.

Merton, R. C. (1974). On the pricing of Corporate Debt: The Risk Strucuture of Interest Rates. *The Journal of Finance*, *29*(2), 449–470.

Newey, W. K., & West, K. D. (1987). A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*, *55*, 703–708.

Rosner, R. L. (2003). Earnings Manipulation in Failing Firms. *Contemporary Accounting Research*, *20*(2), 361–408.

Roychowdhury, S., & Sletten, E. (2012). Voluntary disclosure incentives and earnings informativeness. *Accounting Review*, *87*(5), 1679–1708.

Vassalou, M., & Xing, Y. (2004). Default Risk in Equity Returns. *The Journal of Finance, 59*(2), 831–868.

Welch, I., & Goyal, A. (2008). A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies*, *21*(4), 1455–1508.

Yan, S. (2011). Jump risk, stock returns, and slope of implied volatility smile. *Journal of Financial Economics*, *99*(1), 216–233.

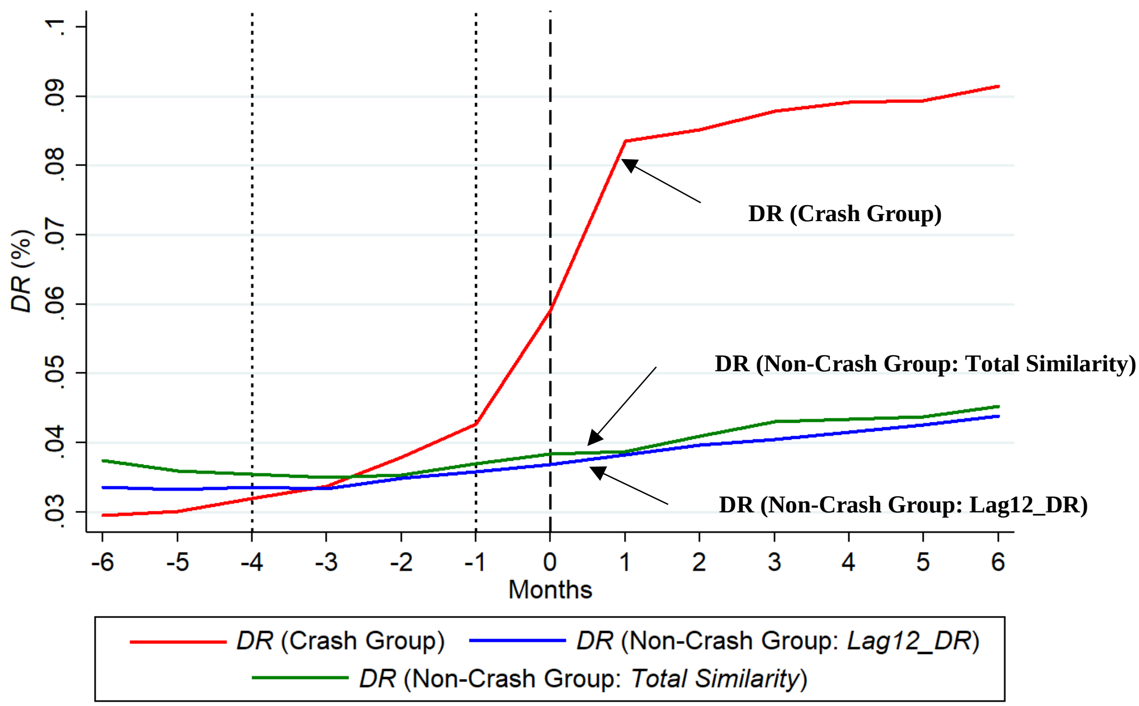
Zhu, W. (2016). Accruals and price crashes. *Review of Accounting Studies*, *21*(2), 349–399.

Figures

Figure 1

Time evolution of distress risk for crashed *vs.* non-crashed firms

The figure illustrates the time evolution of distress risk (*DR*) as per *Eq.* (7) around stock price crashes occurring in month $t = 0$ for: (i) firms that experienced a stock price crash (crashed firms) as depicted by the red line, and (ii) two control-matched samples (non-crashed group). The matching for the samples of non-crashed firms is based on the: (i) total similarity measure of Hoberg and Phillips (2016) matched in month t-1 as depicted by the green line, and (ii) Merton's *DD* value matched in month t-12 as depicted by the blue line.



Tables

Table 1

Summary statistics of yearly and monthly stock price crashes

This table presents in **Panel A** the yearly summary statistics of stock price crashes, short-term changes in distress risk (*ΔDR*) and stock returns (*RET*). Specifically: column (1) features the fiscal years in the sample; column (2) reports the firm observations per year; column (3) reports the number of firm-year crashes using the main stock price crash measure (*CRASH*), whereby a firm-year is flagged as crashed if there is at least one month falling within the fiscal year for which $CRASH_t = 1$; column (4) reports the percentages of crashes per year; column (5) reports the mean return of firms that have crashed within a fiscal year; column (6) reports the mean returns of the firms that have not crashed within a fiscal year; column (7) reports the mean of distress risk changes (*ΔDR*) in month *t* - 1 for firms that have crashed in month *t* within a fiscal year*;* column (8) reports the mean of *ΔDR* in month *t* - 1 for firms that have not crashed in month *t;* column (9) reports the difference between columns (7) and (8). **Panel B** presents an anatomy of monthly stock price crashes across quarterly (QEA) and non-quarterly earnings announcement (non-QEA) months. The quarterly announcement months are defined by using the announcement reported date in Compustat (item "rdq"). **Panel C** reports statistics for stock price crashes and short-term changes in distress risk. Specifically: column (1) features the point in time (quarterly (QEA) and non-quarterly announcement months (non-QEA)); column (2) features the number of monthly observations with stock price crashes; column (3) features the number of observations with no stock price crashes; column (4) reports the mean of distress risk changes (*ΔDR*) in month *t* - 1 for firms that have crashed in month *t*; column (5) reports the mean of *ΔDR* in month *t* - 1 for firms that have not crashed in month *t*; column (6) reports the difference between columns (4) and (5); and column (7) reports the *t*-statistic value for column (6). ***, ** and * indicate statistical significance at the 1, 5 and 10 percent, respectively.

Panel A: Yearly summary statistics for stock price crashes, distress risk changes and stock returns

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Year	**Number of observations**	**Number of crashes**	**Percentage of crashes**	$\overline{RET}_t$ **on crashed**	$\overline{RET}_t$ **for non-crashed**	$\overline{\Delta DR}_{t-1}$ **on crashed [*C*]**	$\overline{\Delta DR}_{t-1}$ **for non-crashed [*NC*]**	**[*C*] - [*NC*]**
1990	788	217	27.54%	-17.24%	-0.10%	2.92%	2.39%	0.53%
1991	859	125	14.55%	-10.12%	3.70%	-0.31%	-2.41%	2.10%***
1992	908	127	13.99%	-17.55%	2.11%	0.86%	0.21%	0.65%*
1993	984	159	16.16%	-15.78%	1.71%	-0.06%	-0.46%	0.41%
1994	1,081	112	10.36%	-17.39%	0.50%	0.34%	0.03%	0.31%
1995	1,178	180	15.28%	-17.64%	2.60%	0.66%	0.15%	0.51%
1996	1,264	193	15.27%	-18.78%	2.11%	0.68%	0.03%	0.65%**
1997	1,352	225	16.64%	-18.69%	1.97%	0.21%	-0.14%	0.35%
1998	1,382	306	22.14%	-16.26%	1.60%	2.36%	1.31%	1.05%***
1999	1,391	242	17.40%	-20.95%	3.58%	1.51%	-0.80%	2.31%***
2000	1,360	295	21.69%	-26.33%	2.01%	2.73%	0.56%	2.16%***
2001	1,385	200	14.44%	-25.10%	3.08%	0.99%	-0.44%	1.43%***
2002	1,515	307	20.26%	-23.95%	-0.56%	2.32%	-0.28%	2.60%***
2003	1,658	236	14.23%	-14.61%	4.78%	-0.59%	-1.92%	1.33%***
2004	1,713	291	16.99%	-17.61%	2.23%	0.51%	-0.31%	0.82%***
2005	1,698	378	22.26%	-16.58%	1.28%	0.45%	-0.19%	0.64%***
2006	1,683	377	22.40%	-17.45%	2.08%	0.51%	-0.21%	0.72%***
2007	1,657	386	23.30%	-17.07%	0.97%	0.64%	-0.01%	0.64%***
2008	1,659	709	42.74%	-20.42%	-2.72%	3.57%	2.77%	0.80%***

2009	1,715	308	17.96%	-19.49%	5.21%	0.60%	-1.22%	1.82%***
2010	1,734	201	11.59%	-15.27%	3.32%	-0.21%	-1.67%	1.46%***
2011	1,665	305	18.32%	-15.89%	0.28%	0.94%	0.28%	0.65%***
2012	1,662	432	25.99%	-14.97%	2.26%	-0.01%	-0.60%	0.59%***
2013	1,685	411	24.39%	-12.13%	3.74%	-0.30%	-0.63%	0.33%*
2014	1,690	438	25.92%	-14.14%	1.03%	0.64%	0.06%	0.58%***
2015	1,703	501	29.42%	-16.30%	0.13%	1.74%	0.73%	1.01%***
2016	1,752	454	25.91%	-16.83%	2.29%	0.94%	-0.52%	1.46%***
2017	1,843	476	25.83%	-16.20%	2.21%	0.64%	-0.61%	1.24%***
2018	1873	419	22.37%	-18.45%	-0.30%	1.10%	-0.27%	1.37%***
Total	42,837	9,010	21.03%	-17.56%	1.91%	1.07%	-0.23%	1.30%***

Panel B: Anatomy of monthly stock price crashes

	(1)	(2)
	Number of crashes	**% of crashes**
Quarterly earnings announcement (QEA) months	5,416	54.23%
Non-quarterly earnings announcement (non-QEA) months	4,572	45.77%
Total occurrences of crashes	9,988	100.00%

Panel C: Stock price crashes and short-term changes in distress risk by QEA and non-QEA months

(1)	(2)	(3)	(4)	(5)	(6)	(7)
Point in time	$CRASH_t = 1$	$CRASH_t = 0$	ΔDR_{t-1} **crash months** ($CRASH_t = 1$)	ΔDR_{t-1} **non-crash months** ($CRASH_t = 0$)	**Difference [column (4) – column (5)]**	***t*-statistic**
Month of QEA 1	1,213	36,257	0.80%	-0.10%	0.90%***	3.64
Month of QEA 2	1,393	36,969	0.40%	-0.18%	0.58%***	-2.64
Month of QEA 3	1,738	37,884	1.01%	0.10%	0.91%***	3.83
Month of QEA 4	1,072	35,830	1.06%	0.06%	1.00%***	3.14
QEA months (total of above)	5,416	146,940	0.81%	-0.03%	0.84%***	6.68
Non-QEA months	4,572	305,750	1.89%	0.00%	1.89%***	10.60
	Difference between non-QEA and QEA months		1.08%***			
	***t*-statistic**		4.98			

Table 2
Summary statistics

This table presents summary statistics for the main variables comprising 462,678 firm-month observations corresponding to 4,855 firms for the period 1990-2018. All variables are defined in the Appendix. Continuous variables are winsorized at the 1st and 99th percentiles. ***, ** and * indicate statistical significance at the 1, 5 and 10 percent, respectively.

Variables	Mean	Std. Dev.	Q1	Median	Q3
Dependent variables					
$CRASH_t$	0.022	0.145	0.000	0.000	0.000
$CRASH\_3.2_t$	0.019	0.136	0.000	0.000	0.000
$CRASH\_20PRC_t$	0.021	0.144	0.000	0.000	0.000
$NCMRET_t$	0.486	0.273	0.310	0.419	0.580
$NCSKEW_t$	0.087	0.991	-0.499	0.042	0.608
$DUVOL_t$	-0.025	0.385	-0.278	-0.037	0.212
Baseline variables					
ΔDR_{t-1}	0.000	0.094	0.000	0.000	0.000
$SIZE_{t-1}$	6.614	1.748	5.276	6.456	7.763
M/B_{t-1}	3.167	4.275	1.395	2.245	3.727
LEV_{t-1}	0.478	0.236	0.301	0.473	0.624
ROA_{t-1}	0.008	0.170	-0.001	0.045	0.085
SA_{t-1}	-3.487	0.577	-3.887	-3.460	-3.095
RET_{t-1}	0.039	0.231	-0.094	0.025	0.149
$DTURN_{t-1}$	0.001	0.078	-0.026	-0.001	0.024
$OPACITY_{t-1}$	0.550	0.435	0.263	0.426	0.684
$NCSKEW_{t-12}$	0.081	0.916	-0.495	0.038	0.595
Additional controls					
$MDLI_{t-1}$	0.004	0.006	0.001	0.002	0.004
CL/CA_{t-1}	0.557	0.383	0.303	0.472	0.702
AGE_10_{t-1}	0.339	0.474	0.000	0.000	1.000
$GOODWILL_{t-1}$	0.086	0.137	0.000	0.000	0.134
HHI_{t-1}	0.291	1.322	0.000	0.001	0.020
$COMPETITIVENESS_{t-1}$	-0.190	1.455	-0.031	0.024	0.099
$R\&D/SALES_{t-1}$	0.187	0.872	0.000	0.000	0.065
$TOBIN'S\_Q_{t-1}$	2.128	1.524	1.214	1.630	2.426
INV_{t-1}	0.025	0.084	-0.009	0.014	0.055

Table 3

Pearson correlation coefficients

This table presents in **Panel A** the Pearson correlation coefficients for the baseline variables and in **Panel B** the Pearson correlation coefficients for the crash risk measures. All variables are defined in the Appendix. Continuous variables are winsorized at the 1st and 99th percentiles. Numbers in bold font type indicate statistical significance at 1%.

Panel A. Correlation coefficients between the baseline variables

	1	2	3	4	5	6	7	8	9	10	11
1. $CRASH_t$	1.000										
2. ΔDR_{t-1}	**0.020**	1.000									
3. $SIZE_{t-1}$	**0.015**	**-0.020**	1.000								
4. M/B_{t-1}	**0.009**	**0.009**	**0.186**	1.000							
5. LEV_{t-1}	0.004	**-0.007**	**0.165**	0.002	1.000						
6. ROA_{t-1}	**0.007**	**0.024**	**0.276**	**-0.025**	**-0.108**	1.000					
7. SA_{t-1}	**-0.011**	-0.003	**-0.652**	**0.071**	**-0.269**	**-0.313**	1.000				
8. RET_{t-1}	**-0.024**	**-0.298**	**0.062**	**-0.014**	**0.014**	**-0.011**	**0.028**	1.000			
9. $DTURN_{t-1}$	**0.012**	**0.044**	**0.034**	**0.059**	**0.028**	**0.036**	**-0.005**	-0.002	1.000		
10. $OPACITY_{t-1}$	0.001	**0.007**	**-0.271**	**0.070**	**-0.054**	**-0.162**	**0.344**	**0.006**	**-0.028**	1.000	
11. $NCSKEW_{t-12}$	**0.010**	**-0.042**	**0.019**	**-0.037**	**0.000**	**-0.039**	**-0.022**	**0.008**	**-0.102**	**0.012**	1.000

Panel B. Correlation coefficient between the crash risk measures

	1	2	3	4	5	6
1. $CRASH_t$	1.000					
2. $CRASH\_3.2_t$	**0.930**	1.000				
3. $CRASH\_20PRC_t$	**0.428**	**0.424**	1.000			
4. $MINRET_t$	**0.257**	**0.256**	**0.310**	1.000		
5. $NCSKEW_t$	**0.178**	**0.179**	**0.131**	**0.435**	1.000	
6. $DUVOL_t$	**0.170**	**0.169**	**0.112**	**0.389**	**0.929**	1.000

Table 4
The impact of short-term changes in distress risk on stock price crashes

This table presents logit regression estimates for the relationship between the 3-month changes in distress risk (ΔDR_{t-1}) and stock price crashes ($CRASH_t$). The estimates feature different regression specifications as per *Eq.* (9). The economic significance of the baseline model (2) is reported in column (3) and is computed as follows: first, the marginal effects are estimated by comparing crash risk at the 25th and 75th percentile values of each variable while holding all other variables at their mean values; second, each variable's economic significance is computed by dividing the marginal effect to the sample mean value of the unconditional probability of stock price crash. Column (4) presents the logit estimates of the baseline model when *CRASH* is recoded to take the value of one when there is a crash that falls in a quarterly earnings announcement (QEA) month; Column (5) presents the logit estimates of the baseline model when *CRASH* is recoded to take the value of one when there is a crash that falls in a non-quarterly earnings announcement (non-QEA) month. The quarterly announcement months are defined by using the announcement reported date in Compustat (item "rdq"). All variables are defined in the Appendix. The regression estimates include a constant, and dummy variables to control for time-invariant year and industry-specific fixed effects. The *z*-statistics are shown in parentheses and are computed based on robust standard errors clustered at the firm-level. All continuous variables are winsorized at the 1st and 99th percentiles and are standardized to have a mean value of zero and variance of one. All the variance inflation factors (VIFs) for the independent variables are less than five, suggesting the absence of any multicollinearity issues. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	$CRASH_t$	$CRASH_t$	**Economic significance of model (2)**	$CRASH\_QEA_t$	$CRASH\_NONQEA_t$	$CRASH_t$	$CRASH_t$	$CRASH_t$	$CRASH_t$
ΔDR_{t-1}	0.111***	0.086***	8.33%	0.064***	0.098***	0.069***	0.083***	0.081***	0.061***
	(11.71)	(8.06)		(4.40)	(6.92)	(6.63)	(7.67)	(7.41)	(5.58)
$SIZE_{t-1}$		0.110***	10.69%	0.091***	0.133***	0.120***	0.097***	0.075***	0.065***
		(7.22)		(4.48)	(6.02)	(7.81)	(5.87)	(4.56)	(3.58)
M/B_{t-1}		0.015	1.48%	0.016	0.011	0.016	0.021**	-0.015	-0.010
		(1.54)		(1.22)	(0.80)	(1.64)	(2.06)	(-1.40)	(-0.88)
LEV_{t-1}		0.022*	2.11%	0.004	0.043**	0.036***	0.023**	0.025**	0.043***
		(1.92)		(0.24)	(2.52)	(2.89)	(1.97)	(2.23)	(3.34)
ROA_{t-1}		0.057***	5.53%	0.087***	0.030*	0.054***	0.082***	0.050***	0.066***
		(4.72)		(4.86)	(1.69)	(4.48)	(5.80)	(4.13)	(4.67)
SA_{t-1}		0.042**	4.12%	-0.014	0.114***	0.047***	0.001	0.005	-0.036
		(2.48)		(-0.58)	(4.47)	(2.73)	(0.06)	(0.28)	(-1.54)
RET_{t-1}		-0.123***	-11.96%	-0.062***	-0.187***	-0.099***	-0.118***	-0.114***	-0.085***
		(-8.84)		(-3.51)	(-8.68)	(-7.26)	(-8.32)	(-8.19)	(-6.08)
$DTURN_{t-1}$		0.059***	5.77%	0.048***	0.061***	0.060***	0.059***	0.056***	0.056***
		(4.37)		(4.15)	(4.26)	(4.41)	(4.32)	(4.40)	(4.40)
$OPACITY_{t-1}$		0.031***	3.01%	0.027*	0.037**	0.032***	0.036***	0.021*	0.027**
		(2.95)		(1.74)	(2.44)	(3.03)	(3.36)	(1.92)	(2.42)
$NCSKEW_{t-12}$		0.067***	6.50%	0.071***	0.060***	0.068***	0.065***	0.070***	0.069***
		(6.80)		(5.35)	(4.21)	(6.87)	(6.54)	(7.11)	(6.89)
$MDLI_{t-1}$						0.279***			0.280***
						(15.58)			(15.24)

CL/CA$_{t-1}$						-0.035***			-0.043***
						(-2.64)			(-3.08)
AGE_10$_{t-1}$							0.076***		0.087***
							(2.65)		(2.93)
GOODWILL$_{t-1}$							0.025**		0.038***
							(2.33)		(3.45)
HHI$_{t-1}$							-0.033**		-0.026*
							(-2.42)		(-1.95)
COMPETITIVENESS$_{t-1}$							-0.052**		-0.049**
							(-2.56)		(-2.57)
R&D/SALES$_{t-1}$							-0.016		-0.026
							(-0.79)		(-1.30)
TOBIN'S_Q$_{t-1}$								0.071***	0.078***
								(5.48)	(5.71)
INV$_{t-1}$								0.065***	0.063***
								(5.16)	(4.80)
YEAR DUMMIES	Yes	Yes	-	Yes	Yes	Yes	Yes	Yes	Yes
INDUSTRY DUMMIES	Yes	Yes	-	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	462,678	462,678	-	462,678	462,678	462,526	445,951	457,532	440,801
Wald Chi-square	1,650.360	1,893.098	-	1,352.593	1,154.245	2,216.923	1,906.131	1,895.667	2,189.090
Log Pseudolikelihood	-47,463.485	-47,307.347	-	-28,795.897	-25,117.477	-47,188.085	-45,514.925	-46,739.980	-44,855.773
Pseudo R^2	0.0151	0.0183	-	0.0230	0.0211	0.0206	0.0187	0.0186	0.0213

Table 5

The impact of distress risk changes on stock price crashes: Alternative distress and crash risk measures

This table presents logit regression estimates for the relationship between 3-month changes in distress risk (ΔDR_{t-1}) and stock price crash ($CRASH_t$) using alternative definitions of the distress and crash risk measures. **Panel A** presents logit regression estimates using as dependent variables the three dichotomous variables, namely *CRASH*, *CRASH_3.2* and *CRASH_20PRC*. **Panel B** reports ordinary least squares estimates using as dependent variables the three continuous-based crash risk measures, namely *NCMRET*, *NCSKEW* and *DUVOL*. The alternative proxy for distress risk is: *ΔDR_ALT* featuring the 3-month changes in distress risk using the model in Charitou *et al*. (2013). All variables are defined in the Appendix. The regression estimates include a constant, and dummy variables to control for time-invariant year and industry-specific fixed effects. All models include the baseline control variables. The *z*-statistics (or *t*-statistics) in Panel A (Panel B) are shown in parentheses and are computed based on robust standard errors clustered at the firm-level (based on Driscoll and Kraay (1998) standard errors). All continuous variables are winsorized at the 1st and 99th percentiles and are standardized to have a mean value of zero and variance of one. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent, respectively.

Panel A: Dichotomous crash risk measures

	(1)	**(2)**	**(3)**	**(4)**	**(4)**	**(5)**
	$CRASH_t$		***$CRASH\_3.2_t$***		***$CRASH\_20PRC_t$***	
ΔDR_{t-1}	0.086***		0.089***		0.114***	
	(8.06)		(7.80)		(11.37)	
ΔDR_ALT_{t-1}		0.090***		0.092***		0.119***
		(8.35)		(7.98)		(11.52)
CONTROLS	Yes	Yes	Yes	Yes	Yes	Yes
YEAR DUMMIES	Yes	Yes	Yes	Yes	Yes	Yes
INDUSTRY DUMMIES	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	462,678	462,678	462,678	462,678	462,678	462,678
Wald chi2	1,893.098	1,903.706	1,750.270	1,763.286	6504.997	6943.475
Log Pseudolikelihood	-47,307.347	-47,305.913	-42,221.296	-42,220.829	-42315.043	-42224.185
Pseudo R^2	0.0183	0.0183	0.0189	0.0189	0.1093	0.1112

Panel B: Continuous crash risk measures

	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
	$NCMRET_t$		***$NCSKEW_t$***		***$DUVOL_t$***	
ΔDR_{t-1}	0.093***		0.037***		0.027***	
	(9.28)		(6.15)		(6.82)	
ΔDR_ALT_{t-1}		0.110***		0.052***		0.038***
		(11.17)		(7.79)		(9.02)
CONTROLS	Yes	Yes	Yes	Yes	Yes	Yes
YEAR DUMMIES	Yes	Yes	Yes	Yes	Yes	Yes
INDUSTRY DUMMIES	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	462,028	462,028	462,678	462,678	462,027	462,027
Adj. R^2	0.190	0.193	0.093	0.094	0.062	0.062

Table 6
Additional analysis based on Merton *DD* component variables

This table presents logit regression estimates for the relationship between 3-month changes in distress risk (ΔDR_{t-1}) and stock price crash ($CRASH_t$) in the presence of additional covariates to control for the influence of variables used in the estimation of distress risk as per *Eq.* (6). Model (1) features the results of baseline model (2) as per Table 4 reported for comparison purposes. The additional variables included are: the stock's 12-month cumulative return measured in month t - 2 (AR_{t-2}); the firm's market-value of assets to book-value of debt (V/D_{t-1}); and the volatility of total assets returns ($\sigma_{BS_{t-1}}$). All variables are defined in the Appendix. The regression estimates include a constant, and dummy variables to control for time-invariant year and industry-specific fixed effects. All models include the baseline control variables. The *z*-statistics are shown in parentheses and are computed based on robust standard errors clustered at the firm-level. All continuous variables are winsorized at the 1st and 99th percentiles and are standardized to have a mean value of zero and variance of one. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent, respectively.

	(1)	(3)	(4)	(5)	(6)
ΔDR_{t-1}	0.086***	0.106***	0.085***	0.082***	0.103***
	(8.06)	(9.66)	(8.06)	(7.75)	(9.35)
AR_{t-2}		0.200***			0.215***
		(18.39)			(18.90)
V/D_{t-1}			0.004		0.002
			(0.67)		(0.33)
$\sigma_{BS_{t-1}}$				-0.132***	-0.160***
				(-8.15)	(-9.67)
CONTROLS	Yes	Yes	Yes	Yes	Yes
YEAR DUMMIES	Yes	Yes	Yes	Yes	Yes
INDUSTRY DUMMIES	Yes	Yes	Yes	Yes	Yes
Obs.	462,678	462,678	462,678	462,678	462,678
Wald Chi-square	1,893.098	2,203.959	1,895.778	1,967.141	2,293.329
Log Pseudolikelihood	-47,307.347	-47,159.754	-47,307.231	-47,274.687	-47,112.923
Pseudo R^2	0.0183	0.0214	0.0183	0.0190	0.0223

Table 7

Time span sensitivity of short-term changes in distress risk on stock price crashes

This table presents logit regression estimates for the relationship between 3-month changes in distress risk measured at various lagged periods (ΔDR_{t-1}, …, ΔDR_{t-6}) and stock price crash ($CRASH_t$). All variables are defined in the Appendix. The regression estimates include a constant, and dummy variables to control for time-invariant year and industry-specific fixed effects. All models include the baseline control variables. The *z*-statistics are shown in parentheses and are computed based on robust standard errors clustered at the firm-level. All continuous variables are winsorized at the 1st and 99th percentiles and are standardized to have a mean value of zero and variance of one. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
ΔDR_{t-1}	0.086***					
	(8.06)					
ΔDR_{t-2}		0.057***				
		(5.54)				
ΔDR_{t-3}			0.051***			
			(5.08)			
ΔDR_{t-4}				0.042***		
				(3.99)		
ΔDR_{t-5}					0.019*	
					(1.85)	
ΔDR_{t-6}						0.010
						(0.92)
CONTROLS	Yes	Yes	Yes	Yes	Yes	Yes
YEAR DUMMIES	Yes	Yes	Yes	Yes	Yes	Yes
INDUSTRY DUMMIES	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	462,678	462,536	462,403	462,364	462,208	462,120
Wald Chi-square	1,893.098	1,838.961	1,833.581	1,817.988	1,802.178	1,790.847
Log Pseudolikelihood	-47,307.347	-47,305.517	-47,305.578	-47,316.613	-47,296.278	-47,291.882
Pseudo R^2	0.0183	0.0179	0.0178	0.0177	0.0176	0.0176

Table 8
Reverse causality tests

Panel A of this table presents the ordinary least squares estimates for the relationship between stock price crash measured at various lagged periods ($CRASH_{t-1}, \ldots, CRASH_{t-6}$) and future changes in distress risk ($\Delta DR_{t,t+3}$). The dependent variable in Panel A, is defined as the short-term change in distress risk from month t to month t+3. **Panel B** of this table presents logit regression estimates for the relationship between 3-month changes in distress risk (ΔDR_{t-1}) and stock price crash ($CRASH_t$) with samples that exclude observations in various periods spanning from one month after the crash as per model (1) up to six months after the crash as per model (6). All variables are defined in the Appendix. The regression estimates include a constant, and dummy variables to control for time-invariant year and industry-specific fixed effects. All models include the baseline control variables. The *t*-statistics (*z*-statistics) in Panel A (Panel B) are shown in parentheses and are computed based on Driscoll and Kraay (1998) standard errors (robust standard errors clustered at the firm-level). All continuous variables are winsorized at the 1st and 99th percentiles and are standardized to have a mean value of zero and variance of one. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent, respectively.

Panel A: Time span sensitivity of stock price crashes on future changes in distress risk

	(1)	(2)	(3)	(4)	(5)	(6)
$CRASH_{t-1}$	0.038**					
	(2.43)					
$CRASH_{t-2}$		0.013				
		(0.88)				
$CRASH_{t-3}$			-0.006			
			(-0.34)			
$CRASH_{t-4}$				0.009		
				(0.62)		
$CRASH_{t-5}$					0.004	
					(0.26)	
$CRASH_{t-6}$						-0.015
						(-0.82)
CONTROLS	Yes	Yes	Yes	Yes	Yes	Yes
YEAR DUMMIES	Yes	Yes	Yes	Yes	Yes	Yes
INDUSTRY DUMMIES	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	458,004	457,935	457,879	457,828	457,817	457,795
R^2	0.029	0.029	0.029	0.029	0.029	0.029

Panel B: The impact of short-term changes in distress risk on stock price crashes: Excluding month-firm observations following the incidence of stock price crashes

	(1)	(2)	(3)	(4)	(5)	(6)
Drop month-firm observations after the crash incidence	**1 month**	**1 & 2 months**	**1, 2 & 3 months**	**1, 2, 3 & 4 months**	**1, 2, 3, 4 & 5 months**	**1, 2, 3, 4, 5 & 6 months**
ΔDR_{t-1}	0.083***	0.085***	0.088***	0.090***	0.092***	0.094***
	(7.51)	(7.26)	(7.18)	(7.12)	(7.30)	(7.48)
CONTROLS	Yes	Yes	Yes	Yes	Yes	Yes
YEAR DUMMIES	Yes	Yes	Yes	Yes	Yes	Yes
INDUSTRY DUMMIES	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	452,946	443,838	435,383	426,952	418,919	411,487
Wald Chi-square	1858.963	1817.045	1774.164	1773.699	1775.196	1775.245
Log Pseudolikelihood	-46313.456	-45282.368	-44124.046	-43510.172	-42874.678	-42077.807
Pseudo R^2	0.0188	0.0191	0.0191	0.0197	0.0202	0.0203

Table 9
Endogeneity tests

This table presents logit regression estimates for the relationship between 3-month changes in distress risk (ΔDR_{t-1}) and stock price crashes ($CRASH_t$) using econometric tests to treat endogeneity. Model (1) is the baseline model reported for comparison reasons. Model (2) augments the baseline model with firm-fixed effect dummies (excluding the industry-fixed effects). Models (3) to (5) report the second stage instrumental variable estimations, whereby the instruments are the monthly-industry-specific small-minus-big (SMB) and high-minus-low (HML) return factors. All variables are defined in the Appendix. The regression estimates include a constant, and dummy variables to control for time-invariant year and industry-specific fixed effects. All models include the baseline control variables. The *z*-statistics are shown in parentheses and are computed based on robust standard errors clustered at the firm-level. All continuous variables are winsorized at the 1st and 99th percentiles and are standardized to have a mean value of zero and variance of one. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent, respectively.

	(1)	(2)	(3)	(4)	(5)
ΔDR_{t-1}	0.086***	0.093***			
	(8.06)	(8.83)			
$\Delta DR\_IV\_SMB_{t-1}$			0.098***		
			(6.56)		
$\Delta DR\_IV\_HML_{t-1}$				0.095***	
				(6.34)	
$\Delta DR\_IV\_HML\&SMB_{t-1}$					0.094***
					(6.32)
CONTROLS	Yes	Yes	Yes	Yes	Yes
FIRM DUMMIES	NO	Yes	No	No	No
YEAR DUMMIES	Yes	Yes	Yes	Yes	Yes
INDUSTRY DUMMIES	Yes	No	Yes	Yes	Yes
Obs.	462,678	413,955	442,213	442,213	442,213
Wald Chi-square	1893.098	1875.295	1796.186	1792.314	1792.325
Log Pseudolikelihood	-47307.347	-40113.017	-45368.851	-45370.457	-45370.545
Pseudo R^2	0.0183	0.0228	0.0180	0.0179	0.0179

Table 10
The Sarbanes–Oxley Act as a quasi-experimental setting

Panel A of this table presents logit regression estimates investigating the effect of quarterly earnings announcement months ($DQEA_t$) falling in the post Sarbanes–Oxley period ($DPOST\_SOX_t$) and stock price crashes ($CRASH_t$). The sample spans the period from July 1999 to June 2005. *DPOST_SOX* is a binary variable that takes the value of one in the period from July 2002 to June 2005 (POST-SOX), and zero in the period from July 1999 to June 2002 (PRE-SOX). *DQEA* takes the value of one for months that a firm is making a quarterly earnings announcement, and zero otherwise. The quarterly announcement months are defined by using the announcement reported date in Compustat (item "rdq"). All variables are defined in the Appendix. The regression estimates include a constant, and dummy variables to control for time-invariant year and industry-specific fixed effects. All models include the baseline control variables. The *z*-statistics are shown in parentheses and are computed based on robust standard errors clustered at the firm-level. All continuous variables are winsorized at the 1st and 99th percentiles and are standardized to have a mean value of zero and variance of one. **Panel B** of this table tabulates the mean values of the 3-month changes in distress risk (ΔDR_{t-1}) for the quarterly earnings announcement (QEA) months and the non-quarterly earnings announcement (non-QEA) months, across cases when a crash occurs ($CRASH_t = 1$) *vs.* cases with no crash occurring ($CRASH_t = 0$). ***, **, and * indicate statistical significance at the 1, 5, and 10 percent, respectively.

Panel A: Logit regression estimates	
	$\boldsymbol{CRASH_t}$
$\boldsymbol{DPOST\_SOX_t}$	0.161
	(1.28)
$\boldsymbol{DQEA_t}$	0.300***
	(3.99)
$\boldsymbol{DPOST\_SOX_t \times DQEA_{t_t}}$	0.652***
	(6.57)
CONTROLS	Yes
YEAR DUMMIES	Yes
INDUSTRY DUMMIES	Yes
Obs.	100,107
Wald Chi-square	489.141
Log Pseudolikelihood	-8,652.218
Pseudo R^2	0.0279

Panel B: Mean values of the 3-month changes in distress risk measured in month *t*-1 ($\boldsymbol{\Delta DR_{t-1}}$) aggregated, in different periods (QEA *vs.* non-QEA), and across crash *vs.* non-crash events measured in month *t*

Quarterly earnings announcement (QEA) months				
	(1)	**(2)**	**(3)**	**(4)**
	PRE-SOX	**POST-SOX**	**Difference**	***t*-statistic**
$\boldsymbol{CRASH_t = 1}$	-0.58%	0.75%	1.32%**	2.15
$\boldsymbol{CRASH_t = 0}$	-0.35%	-0.19%	0.16%	1.42

Non-quarterly earnings announcement (non-QEA) months				
	PRE-SOX	**POST-SOX**	**Difference**	***t*-statistic**
$\boldsymbol{CRASH_t = 1}$	0.66%	1.36%	0.70%	1.01
$\boldsymbol{CRASH_t = 0}$	-0.23%	-0.28%	-0.04%	-0.55

Table 11

The impact of short-term changes in distress risk on stock price crashes: The hoarding of bad news effect

Panel A of this table presents the recursive regression estimates of *Eq.* (11) investigating the impact of short-term changes in distress risk (ΔDR_{t-1}) on future financial reporting opacity. The dependent variable, $\Delta OPACITY\_3M_t$, is the change in accounting opacity from month *t*-3 to *t*. The *t*-statistic is shown in parenthesis and is computed based on Newey and West (1987) standard errors. The **Panel B** of this table presents logit regression estimates for the relationship between the estimated financial opacity attributed to the short-term changes in distress risk (*i.e.*, $\Delta DR\_\widehat{OPACITY}\_3M_{t-1}$) and stock price crashes ($CRASH_t$). All variables are defined in the Appendix. The regression estimates include a constant, and dummy variables to control for time-invariant year and industry-specific fixed effects. All models include the baseline control variables. The *z*-statistic is shown in parenthesis and is computed based on robust standard errors clustered at the firm-level. All continuous variables are winsorized at the 1st and 99th percentiles and are standardized to have a mean value of zero and variance of one. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent, respectively.

Panel A: Regression estimates of the impact of short-term changes in distress risk measured in month *t*-1 (ΔDR_{t-1}) on changes in financial reporting opacity measured in month *t*		**Panel B: Logit regression estimates**	
	$\Delta OPACITY\_3M_t$		$CRASH_t$
ΔDR_{t-1}	0.018***	$\Delta DR\_\widehat{OPACITY}\_3M_{t-1}$	0.072***
	(13.18)		(5.89)
CONTROLS	Yes	***CONTROLS***	Yes
YEAR DUMMIES	Yes	***YEAR DUMMIES***	Yes
INDUSTRY DUMMIES	Yes	***INDUSTRY DUMMIES***	Yes
		Obs.	318,733
		Wald Chi-square	1,286.742
		Log Pseudolikelihood	-32,548.006
		Pseudo R^2	0.0177

Table 12

The impact of short-term changes in distress risk on stock price crashes: Information asymmetry

This table presents logit regression estimates for the relationship between 3-month distress risk changes ($\Delta DR_{i,t-1}$) and stock price crashes ($CRASH_t$) under various conditions of information asymmetry based on analyst's coverage and analysts' dispersion (*AFD*). Analysts' dispersion is defined as the standard deviation of analysts' earnings per share forecasts divided by the absolute value of the average analysts' forecasts. Models (1) to (3) present the stocks sorted based on analyst coverage (number of analysts following the firm) in month *t*, where model (1) includes the stocks with the lowest or without analyst coverage (firms with missing analyst data are recorded as zero coverage and included in the bottom tercile), while model (3) includes the stocks with the highest analyst coverages. Models (4) and (5) include the stocks sorted based on analysts' forecast dispersion in month *t*. All variables are defined in the Appendix. The regression estimates include a constant, and dummy variables to control for time-invariant year and industry-specific fixed effects. All models include the baseline control variables. The *z*-statistics are shown in parentheses and are computed based on robust standard errors clustered at the firm-level. All continuous variables are winsorized at the 1st and 99th percentiles and are standardized to have a mean value of zero and variance of one. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent, respectively.

	(1)	**(2)**	**(3)**	**(4)**	**(5)**
	Analysts coverage			**Analysts' earnings forecast dispersion**	
	Bottom tercile	**Middle tercile**	**Top tercile**	**Below median *AFD***	**Above median *AFD***
Level of information asymmetry:	**High**	**Medium**	**Low**	**Low**	**High**
ΔDR_{t-1}	0.122***	0.069***	0.040	0.002	0.082***
	(7.87)	(3.82)	(1.63)	(0.07)	(5.96)
$SIZE_{t-1}$	0.265***	0.119***	0.050*	0.064**	0.118***
	(7.50)	(3.25)	(1.65)	(2.55)	(4.80)
M/B_{t-1}	-0.009	0.030*	0.016	0.021	0.028**
	(-0.43)	(1.78)	(0.98)	(1.39)	(2.01)
LEV_{t-1}	0.053***	0.008	0.007	-0.007	0.025
	(2.81)	(0.43)	(0.31)	(-0.37)	(1.44)
ROA_{t-1}	0.048**	0.060***	0.025	0.003	0.082***
	(2.54)	(2.91)	(0.84)	(0.10)	(4.40)
SA_{t-1}	0.044	0.028	0.084***	0.033	0.040
	(1.49)	(1.00)	(2.74)	(1.28)	(1.56)
RET_{t-1}	-0.028	-0.144***	-0.250***	-0.195***	-0.140***
	(-1.25)	(-6.59)	(-9.61)	(-8.15)	(-7.46)
$DTURN_{t-1}$	0.052***	0.065***	0.069***	0.073***	0.069***
	(3.36)	(4.52)	(3.94)	(4.23)	(6.75)
$OPACITY_{t-1}$	0.038**	0.038**	0.006	0.036*	0.029*
	(2.15)	(2.21)	(0.30)	(1.86)	(1.79)
$NCSKEW_{t-12}$	0.081***	0.071***	0.022	0.039**	0.071***
	(4.73)	(4.56)	(1.22)	(2.44)	(4.93)
YEAR DUMMIES	Yes	Yes	Yes	Yes	Yes
INDUSTRY DUMMIES	Yes	Yes	Yes	Yes	Yes
Obs.	174,531	144,068	144,079	178,275	175,842
Wald Chi-square	834.710	783.036	878.443	724.963	1102.663
Log Pseudolikelihood	-15,909.594	-16,109.022	-15,070.705	-18,871.991	-19,276.935
Pseudo R^2	0.0235	0.0196	0.0219	0.0163	0.0241

Table 13

The behavior of short-term changes in distress risk and stock price crashes during analysts' recommendation revisions

This table presents the behavior of the 3-month changes in distress risk (ΔDR_{t-1}) across crashed ($CRASH_t$ = 1) and non-crashed ($CRASH_t$ = 0) months during analysts' recommendation revisions. **Panel A** focuses on increases in analysts' "sell" recommendations, while **Panel B** focuses on decreases in analysts' "buy" recommendations. Both panels present the analyses based on three samples: (i) full sample for which there is an analyst recommendation revision (top part), (ii) the sample for which an analyst recommendation revision occurs within a quarterly earnings announcement (QEA) month (middle part), and (iii) the sample for which an analyst recommendation revision occurs within a non-quarterly earnings announcement (non-QEA) month (bottom part). An increase in a "sell" recommendation is defined when the percentage of analysts with sell recommendations increase from month t - 1 to month t. A decrease in "buy" recommendations is defined when the percentage of analysts with buy recommendations is decreased from month t - 1 to month t. For each case, a comparison analysis based on a control-matched sample is provided. The matched samples are defined based on firms' size for each month and industry. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent, respectively.

Panel A: Increase in Sell Recommendations					**Panel B: Decrease in Buy Recommendations**				
	(1)	**(2)**	**(3)**	**(4)**		**(1)**	**(2)**	**(3)**	**(4)**
		Full Sample					**Full Sample**		
	***CRASH$_t$* = 1**	***CRASH$_t$* = 0**	**Difference**	***t*-statistic**		***CRASH$_t$* = 1**	***CRASH$_t$* = 0**	**Difference**	***t*-statistic**
ΔDR_{t-1}	2.56%	0.55%	2.01%***	4.56	ΔDR_{t-1}	1.43%	0.47%	0.96%***	4.83
Obs.	1,015	25,760			**Obs.**	2,989	66,369		
		Matched sample					**Matched sample**		
ΔDR_{t-1}	2.56%	0.53%	2.03%***	3.75	ΔDR_{t-1}	1.43%	0.37%	1.06%***	3.39
Obs.	1,015	369			**Obs.**	2,989	990		
		QEA months					**QEA months**		
	***CRASH$_t$* = 1**	***CRASH$_t$* = 0**	**Difference**	***t*-statistic**		***CRASH$_t$* = 1**	***CRASH$_t$* = 0**	**Difference**	***t*-statistic**
ΔDR_{t-1}	1.64%	0.58%	1.06%**	2.20	ΔDR_{t-1}	0.93%	0.48%	0.45%**	1.97
Obs.	547	8,651			**Obs.**	1,641	22,551		
		Matched sample					**Matched sample**		
ΔDR_{t-1}	1.64%	0.64%	1.00%	1.39	ΔDR_{t-1}	0.93%	0.18%	0.75%**	1.98
Obs.	547	151			**Obs.**	1,641	383		
		Non-QEA months					**Non-QEA months**		
	***CRASH$_t$* = 1**	***CRASH$_t$* = 0**	**Difference**	***t*-statistic**		***CRASH$_t$* = 1**	***CRASH$_t$* = 0**	**Difference**	***t*-statistic**
ΔDR_{t-1}	3.64%	0.54%	3.10%***	4.21	ΔDR_{t-1}	2.05%	0.47%	1.58%***	4.59
Obs.	468	17,109			**Obs.**	1,348	43,818		
		Matched sample					**Matched sample**		
ΔDR_{t-1}	3.64%	0.46%	3.18%***	3.70	ΔDR_{t-1}	2.05%	0.50%	1.55%***	3.19
Obs.	468	218			**Obs.**	1,348	607		