Bank Financial distress and stock price crashes

Priti Biswas[1], Debasish Maitra[21]

[1]Doctoral student at Indian Institute of Management Indore, India, email: f21pritib@iimidr.ac.in

[2]Indian Indian Institute of Management Indore, email: debasishm@iimidr.ac.in

Abstract

Using 1208 unique listings comprising 118,292 bank-month observations from 2000 to 2022, we examine the effects of short-term changes in banking firms' financial distress risk or probability of default of stock price crash risk. There are significant impacts of short changes in the distress on banks' stock price crash risk. The study also employs different alternative measures of crash risks and distress risk and documents the consistency in the findings. We also find that the channel through which the distress risk reflects on the stock price crash risk is bank opacity as proxied by loan loss provisions. Contrary to similar literature on non-financial firms, we find a negative association between opacity conditional on increasing default risk and future stock price crash risk**.** Moreover, we identify banks with high and low distress risk, which are similar across other covariates. The study provides light for investors who need to know the sensitivity of stock prices to the distress risk of banking firms.

Keywords: Distress risk, crash risk, opacity, loan loss provisions

GEL Classification: G11, G12, G19

[1] Corresponding Author

Bank Financial distress and stock price crashes

1. Introduction

The existing literature has investigated the managerial tendency to withhold critical information about the firm and their subsequent impact on firm stock prices (Andreou et al., 2016, 2017, 2021; Callen and Fang, 2013; Jin and Myers, 2006; Kim et al., 2011). When firm management hides bad news for a prolonged period of time, the firm value appears inflated as investors perceive the firm to be worth more than it truly is. However, firm management cannot keep such news hidden beyond a point when all the accumulated news is revealed at once. As a result, investors penalize the firms, and stock prices crash because of the associated fear. Banking firms are no exception to this. The global financial crisis (GFC) of 2008 showed that banks tend to hide or hoard bad news, and stock prices tumble when the news is finally revealed to the market. Sometimes, the hoarding of bad news by managers in banking firms causes stock price crashes; at other times, this leads to worse outcomes like instability in the financial system, emphasizing the importance of bank-specific news (Barth et al., 2012). Extant literature on bad news hoarding in banks and subsequent stock price crashes show that this effect is propagated through bank loan cycles and non-performing loan (NPLs) management (Andreou et al., 2017; Cohen et al., 2014), manipulation or underreporting of loan loss provisions (or opacity) through different accounting practices (Bushman and Williams, 2015, Jung et al., 2019), and risk management instruments employed by banks (Dewally and Shao, 2013).

It is evident the banks tend to adopt practices to delay or underrecognize the credit losses associated with loans and credit instruments (FASB, 2012). Such behaviour may be especially triggered when the bank is facing sudden financial distress. Bank managers have considerable discretion over the amount of loan loss provisions (*LLPs*) which are to be reported. There are two channels through which a bank's underreporting or overreporting of the actual amount of loan losses may be associated with stock price crashes. On one hand, high *LLPs* could signal to investors that the bank has higher non-performing loans on the balance sheet, which may affect the bank's financial stability. Banks with higher NPLs are more likely to default; thus, investors' reaction to higher LLPs may cause a stock price crash. On the other hand, lower *LLPs* may be a result of withholding bad news of credit losses for a long period of time, which, when revealed, leads investors to react to the simultaneous release of news abruptly and, thus, results in a sharp fall in stock prices or stock price crash.

The argument by Andreou et al. (2017) is that most of the distress risk and its association with future stock price crash risk are understood through panel regressions wherein the accounting-based measures of default risk are employed at annual frequency; thus, it may not be able to fully capture the effects of default risk on futures stock price crash risk as firms may hold up the bad news for longer period of time, up to three or more years (Andreou et al., 2017 and 2021; Bleck and Liu, 2007; Jin and Myers, 2006). However, investors discover some part of the bad news being hidden, and they then react to the stock prices by revising their expectations about the firm's stability. The discovery of bad news by investors occurs before managers' actual disclosure of such news (Andreou et al., 2021). This results in a temporary, short-term increase in the firm's default risk before the stock price actually crashes. Thus, literature focused on distress risk measured at annual frequency may not be able to capture the true relationship between distress risk and stock price crash risk because the association is much more short-term. It is evident in the argument and findings of Andreou et al. (2021) and Chava and Jarrow (2004) that distress risk must be measured at a monthly frequency.

The previous studies have employed different rating models, market-based or accounting measure-based *Z* score, logistic regression-based probability of default, and distance to default (*DTD*) by Merton (1974), as measures of distress risk. Merton's (1974) *DTD,* in particular, employs market-based variables that factor in investors' expectations about the firm to estimate distress risk.

Armed with the above arguments, we examine the relationship between distress risk and stock price crash risk for banking firms in the US. Although bank managers exhibit a tendency to withhold news about the bank's growing financial distress, the bad news may not be kept hidden for a long period as market analysts continuously track the bank's stock and performance; thus, the information is ultimately revealed to the market in the short run. Based on the information revealed, investors revise the price of the shares. Following the strategy of Andreou et al. (2015, 2021), we employ short-term changes in the distress risk and examine the effects of such changes on the stock price crash risk. To the best of our knowledge, the extant literature has not given much emphasis on the relationship between distress risk and stock price crash risk for banking firms. Banking firms are different from non-financial firms as banks are subject to different regulatory frameworks and stricter capital requirements. As financial intermediaries, banking firms are more vulnerable to different sources of risk, such as credit and liquidity risks (Schuermann and Strioh, 2002), and thus, are required to adopt different accounting treatments to safeguard against such risks. In particular, banks are required to maintain adequate provisions to buffer against future expected loan losses. Banking firms employ different measures to forecast loan losses and accordingly decide upon the amount of loan loss provisions. Such provisions are, thus, subject to managerial choice and discretion (Hegde and Kozlowski, 2021). Higher discretion over reported LLPs may result in higher or lower-than-expected reported LLPs, which reduces the predictive power of LLP models. This results in less transparent financial statements and, thus, higher opacity. For non-financial firms, higher opacity, as measured through the amount of discretionary accruals (Dechow et al. 1995; Jones et al. 1991), has been found to be positively associated with stock price crash risk (for instance, Hutton et al. 2009). For banking firms, the association is more ambiguous. A high level of loan loss provisions may indicate the bank's expectations of higher future default risk; however, higher loan loss provisions, especially the discretionary component, have shown a positive effect on stock returns (Beaver and Engel, 1996; Hegde and Kozlowski, 2021; Kilic et al., 2013; Liu et al., 1997). This is contrary to non-banking firms. Despite banking firms' unique yet important nature, the literature has not provided much evidence about the relationship between short-term changes in distress risk and stock price crash risks.

We consider 1208 unique listed banks and bank holding companies incorporated in the US over the period January 2000 to December 2022, comprising 118,292 bank-month observations. Bank financial distress risk is measured using the naïve distance-to-default and probability of default (*PD*) approach first proposed by Bharath and Shumway (2008). We examine the effect of short-term (3-month) change in the probability of default. Our primary crash measure is a binary variable, which takes a value of 1 for the month when at least one firm-specific weekly return of the bank is 3.09 standard deviations lower than the average firm-specific weekly return over the estimation period. We find a significant positive effect of a three-month increase in the probability of default on future crash risk after controlling relevant bank or firm-specific variables (Andreou et al., 2015, 2017, 2021; Beaver and Engel, 1996; Doan et al., 2024; Jung et al., 2019; Kim et al., 2011).

To scrutinize the relationship between short-term changes in the distress risk and stock price crash of banking firms, we conduct a series of tests with alternative measures of distress risk and stock price crash risk. We apply distance-to-capital (Chan-Lau and Sy, 2009) as an alternative measure of changes in distress risk. This measure was curated by Chan-Lau and Sy (2009), especially for banking firms. Given that the Basel Committee on Banking Supervision recommends capital requirements for banks and is strictly monitored by regulators, the distance-to-capital adjusts distress risk with capital requirements under the BASEL regime. We also employ three alternative measures of crash risk, including two continuous measures. Our results are also robust against a series of tests for omitted variable bias and reverse causality. In addition, to ensure our results are not affected by bias arising from functional-form misspecification, we employ propensity score matching (PSM) to identify banks with high and low distress risk which are similar across other covariates. Our results continue to hold for the propensity score-matched sample.

Having established a positive association between short-term increases in default risk and stock price crash risk, and consistent with the arguments in the literature, we delve deeper into finding the channel through which managers in banks try to hide the bad news. We first follow the literature on bank opacity (Zheng, 2020; Jiang et al., 2016) and estimate bank opacity as the absolute discretionary LLPs (DLLPs). Contrary to similar literature on non-financial firms, we find a negative association between opacity conditional on increasing default risk and future stock price crash risk, suggesting that higher opacity reduces crash risk and, thus, higher earnings management. To investigate this apparent aberration, we then employ *raw DLLPs* as a measure of bank opacity to take into consideration that for banking firms specifically, the sign, and not just the magnitude of DLLPs, are indicative of opacity. We find that short-term increases in financial distress are negatively associated with DLLPs, which may indicate that banks tend to report *lower* LLPs in the face of increasing distress. Further, we find that DLLPs are negatively associated with future stock price crash risk. This indicates that banks that hold lower LLPs in response to rising distress have higher stock price crash risks. On the other hand, banks which build up provisions (through higher DLLPs), are protected against future financial shocks and are thus associated with lower crash risk.

Our study contributes to the existing literature in three major ways. First, our high-frequency or monthly measures of distress risk capture more information and reveal a positive association of short-term increase in default risk on future stock price crash risk. Second, consistent with the evidence in Andreou et al. (2021), this study documents that bad news hoarding by bank management is a connect between distress risk and stock price crash risk. Contrary to the findings of non-banking firms and in tune with the literature on banking, we notice that discretionary loan loss provisions as a proxy of a bank's opacity show a negative relationship with crash risk. This is because a higher amount of provisioning- under the paramount pressure of regulators- provides confidence to the investors that banks follow conservative practices to build up a higher buffer to strengthen the balance sheet against probable financial adversities in the future. Finally, this study lends support to extant literature on crash risk and finds that investors of banking firms are no different, and they constantly monitor the riskiness of banks through different sources without being dependent on financial results published by the banks.

The rest of the paper is organized as follows: Section 2 motivates our study through case studies of recent bank defaults. Section 3 describes the data, variables, and methodology. Section 4 highlights the preliminary results. Section 5 provides a full-length analysis of the findings. Section 6 presents robustness checks. Section 7 concludes the paper.

2. Case study of bank failures

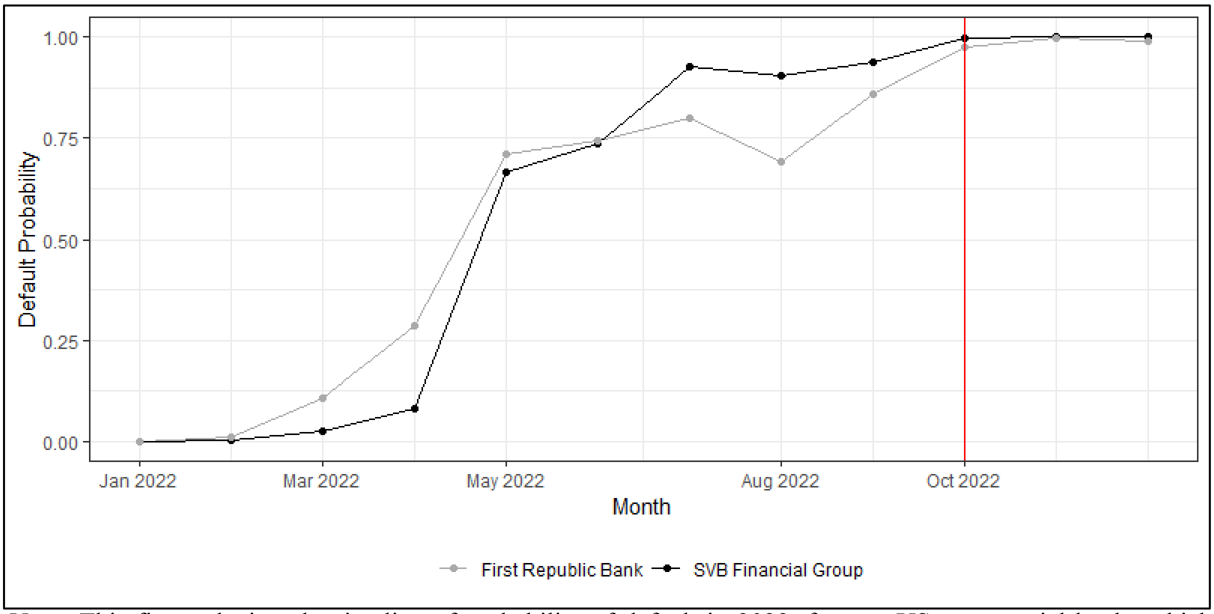
We examine two recent instances of bank distress and failure. The year 2023 saw the demise of a number of banks in the United States, notable among these being the Silicon Valley Bank and the First Republic Bank. The Silicon Valley Bank was a major regional bank which catered mostly to technology start-ups. Its growth was monumental in the pandemic of 2020-2021, fuelled in large part by the growth in its clients in the technology sector[2]. The bank had invested a major chunk of deposits in long-term treasury bonds, whose market value began to take a hit when the regulator hiked market interest rates in an attempt to curb inflationary pressures. Following the declining value of investments and an increase in cash withdrawal by their clients, the bank had to sell off part of their investments at considerable loss, a news which was revealed in March 2023. The ensuing bank run forced the FDIC to step in and take over the bank, resulting in what is being termed the biggest banking collapse since the GFC. In an attempt to examine the financial distress prior to their collapse, we plot the default probabilities of SVB Financial, the holding company of Silicon Valley Bank, in Fig.1.

We also plot the same for First Republic Bank (Fig.1), another banking institution that collapsed during the same time as the Silicon Valley Bank. We observe that, while default probability was low even at the beginning of 2022, it has risen rapidly since March 2022, ultimately culminating in a stock price crash in October 2022. The rapid increase in distress was evident for both banks as far as six months before the stock price crash, and almost four months before the ultimate collapse of both banks in March 2023. Interestingly, a report by the Federal Reserve post-collapse of Silicon Valley Bank attributed the collapse partly to a failure by bank management to manage its liquidity position and a failure of its board to oversee crucial

[2] For more details, the reader is referred to: https://www.nytimes.com/article/svb-silicon-valley-bank-explainer.html

management decisions[3]. A bank run on Silicon Valley Bank in March 2023 was ultimately triggered by its announcement of unreported investment losses, which led to its crash and subsequent collapse.

Fig 1: Timeline of Probability of Default of SVB Financial and First Republic Bank



Note: This figure depicts the timeline of probability of default in 2022, for two US commercial banks which defaulted in 2023. The plot in black depicts the evolution of default probability for SVB Financial Group, the holding company for Silicon Valley Bank that defaulted in March, 2023, and was acquired by First Citizens Bank. The plot in grey depicts the default probability of First Republic Bank, another US commercial bank that defaulted and was seized by the FDIC in May, 2023, and was subsequently acquired by JP Morgan Chase and Co. The vertical red line indicates the month in which there was a stock price crash for the banks (October 2022).

3. Data, variables, and empirical strategy

3.1. Data

Our dataset comprises banks and bank holding companies incorporated in the US. We screen all companies having industry types as banks in the S&P Global Capital IQ database. We restrict our sample to banks incorporated in the US. We merge this data with Compustat North America.[4] Data on daily stock prices come from Compustat daily files. We screen the data to keep only common shares[5] being traded in USD. We drop banks having less than twelve

[3] The detailed report by the FED is available at https://www.federalreserve.gov/newsevents/pressreleases.htm
[4] Common Identification Key (CIK) is the preferred key used for getting the data and merging across databases as it is not reused, unlike tickers. However, CIKs may change over the life of a company; for instance, when the company undergoes an acquisition or a change in legal name. Under these circumstances, a single company may have multiple CIKs. We manually merge banks which have multiple CIKs in the S&P Global Capital IQ database.
[5] Compustat item tpci=0 identifies common shares.

months of daily stock price data. Data on quarterly gross loans, net loans, total deposits, earnings before tax (EBT), loan-to-deposit ratio, and non-performing loans ratio come from the S&P Capital IQ database. Data on other quarterly fundamentals come from Compustat Quarterly files. While merging market-level data with bank-specific fundamentals, we ensure that monthly market-level data is merged with one-quarter-lagged fundamentals, keeping in mind that quarterly fundamentals were made available to the public with a lag. After merging data and removing banks with missing data, our final sample comprises 1208 unique banks[6] over the period January 2000 to December 2022.

3.2. Variables

3.2.1. Crash risk measures

Following the stock price crash risk literature (Andreou et al. 2021; Kim et al., 2011), we estimate multiple measures for crash risk. As we employ a bank-month panel, each crash measure is estimated monthly. We first estimate firm-specific weekly returns using the following market model:

$$r_i^w = \beta_1 + \beta_{2,i} r_m^{w-2} + \beta_{3,i} r_m^{w-1} + \beta_{4,i} r_m^w + \beta_{5,i} r_m^{w+1} + \beta_{6,i} r_m^{w+2} + \epsilon_i^w \quad (1)$$

Here, r_i^w is the weekly return of firm *i* in week *w,* and r_i^w is the weekly return on the market index in week *w.* Following prior literature, and to ensure crash risk measures are relevant to each specific time period, Eq. (1) is estimated on a rolling window of the most recently available 52-weeks of data, where the month in question is the last month within the estimation window. Firm-specific weekly returns are then calculated as the natural log of the residuals from Eq. (1) plus 1:

$$R_i^w = \ln(1 + \epsilon_i^w) \quad (2)$$

Our first and primary crash measure is *c309*, a binary variable that takes a value of 1 for the month when at least one firm-specific weekly return of the bank is 3.09 standard deviations lower than the average firm-specific weekly return over the estimation period. We also employ another crash measure, *c320*, a binary variable that takes a value of 1 for the month when at least one firm-specific weekly return of the bank is 3.20 standard deviations lower than the average over the estimation period. We also consider two continuous crash risk measures

[6] The final sample covers GICS Sub-industries 40101010 (Diversified Banks), 40101015 (Regional Banks), 40102010 (Thrifts and Mortgage Finance), and 40202010 (Consumer Finance). https://www.msci.com/our-solutions/indexes/gics

for our robustness tests: negative conditional skewness (NCSKEW) and down-to-up volatility (*DUVOL*). *NCSKEW* is the negative of the third moment of firm-specific weekly returns, scaled by the cube of their standard deviation. *DUVOL* is the natural logarithm of the ratio of the standard deviation of "up" weeks to that of "down" weeks, where "up" and "down" weeks are when the firm-specific weekly returns fall above and below the mean over the estimation window. These measures are formally estimated using Eqs.(3) and (4):

$$NCSKEW_i^t = \frac{-n(n-1)^{\left(\frac{3}{2}\right)}\sum(R_i^w)^3}{(n-1)(n-2)(\sum(R_i^w)^3)^{\left(\frac{3}{2}\right)}} \tag{3}$$

$$DUVOL_i^t = \ln(\frac{(n_u-1)\sum_{Down}(R_i^w)^2}{(n_d-1)\sum_{Up}(R_i^w)^2} \tag{4}$$

3.2.2. Measures of Financial Distress

We measure financial distress using the naïve distance-to-default measure first proposed by Bharath and Shumway (2008). This measure builds upon the original Merton's distance-to-default model by proposing estimation using company fundamentals, which is simpler than the original Merton's DD model that required one to solve a system of simultaneous non-linear equations. We first estimate the monthly distance-to-default (DTD) using Eq.(5) below:

$$DTD_i^t = \frac{\ln\left(\frac{FV}{D}\right) + (AR_i^{t-1} - (0.5)\sigma_{AS}^2)(T)}{\sigma_{AS}\sqrt{T}} \tag{5}$$

Here, *FV* is the total firm value, which is the sum of the market value of equity and the book value of debt (*D*). Following Nagel and Purnanandam (2020), we take the book value of debt as the sum of total debt, total deposits, and total preferred equity. This differs from the methodology followed for non-financial firms, where the book value of debt is calculated as the sum of short-term debt and half of long-term debt. As Nagel and Purnanandam (2020) argue, this approach may not adequately capture the default risk of banks for two reasons. First, external deposits form a major portion of banks' outstanding liabilities, most of which are short-term. Second, the bank may not be able to roll over such short-term debt if it fails to meet its obligations on long-term debt. We follow their convention in our study when calculating the outstanding debt. *AR* is the cumulative equity returns of the bank over the prior 12-month period. σ_{AS} is asset volatility, estimated as the weighted average of equity and debt volatility, while *T* approximates debt maturity, which is 1 year. Having estimated the *DTD*, probability of default (*PD*) is estimated using Eq.(6):

$$PD_i^t = N(-DTD_i^t) \tag{6}$$

Finally, the 3-month change in probability of default is estimated using Eq.(7):

$$\Delta PD_i^t = PD_i^t - PD_i^{t-3} \tag{7}$$

3.2.3. Measures of bank opacity

The theoretical underpinnings of opacity are rooted in the earnings management literature, which employs discretionary accruals as a proxy for the extent of managerial earnings manipulation (Dechow et al., 1995; Healy, 1985; Jones, 1991). For banks, loan loss provisions (*LLPs*) are argued to be the major vessel for managing reported earnings and disclosures (Dechow et al. 2010). The unexplained or *abnormal* portion of *LLPs* from a predictive model of *LLPs* gives an indication of the extent of *discretion* exercised by bank management over reported *LLPs*. A larger unexplained portion of *LLPs* indicates more opaque financial disclosures. Thus, opacity effectively proxies for the quality of disclosure in financial statements. If discretionary *LLPs* are relatively low, then financial statements are expected to reflect the bank's true state of financial health. However, in the presence of considerable managerial discretion over reported *LLPs*, financial statements lose their informational effectiveness and, in essence, become more opaque to external stakeholders. Based on the literature, we employ two models to predict discretionary LLPs. We first follow Jiang et al. (2016) and use the following regression specification with bank-fixed effects for predicting LLPs:

$$\begin{aligned} LLP_i^t &= \alpha_0 + \alpha_1 dNPL_i^{t-1} + \alpha_2 dNPL_i^t + \alpha_3 dNPL_i^{t+1} + \alpha_4 Size_i^{t-1} \\ &+ \alpha_5 loang_i^t + \alpha_6 dGDP^t + \alpha_7 CSHPI^t + \alpha_8 dUNEMP^t + \theta_i + \epsilon_i^t \end{aligned} \tag{8}$$

Here, LLP_i^t is the *LLP* of bank *i* in quarter *t*, $dNPL_i^t$ is the quarter-on-quarter change in non-performing loans to total loans ratio, *Size* is the natural log of total assets, loang is the quarter-on-quarter loan growth. Both concurrent as well as one-quarter lagged and leading non-performing loans are included in the model because non-performing loans over all three periods may be considered by bank management while deciding upon the actual amount of *LLPs* (Bushman and Williams, 2012). The model also employs three macro-economic variables to control for the possible effects of the external economic environment on *LLPs*: *dGDP*, which is the change in the national gross domestic product; *CSHPI,* the Case-Shiller House Price Index[7]; and change in unemployment rate *dUNEMP*. The first measure of opacity, *Opacity1*, is then estimated as the natural log of absolute residuals from Eq. (8). Absolute residuals, and not positive/negative residuals, are considered to account for the fact that the magnitude, and not the sign, of residuals, is indicative of discretionary *LLPs* (*DLLPs*). For the second measure of opacity, we estimate the following predictive model with bank and quarter fixed effects for *LLPs* as in Zheng (2020):

$$\begin{aligned} LLP_i^t &= \alpha_0 + \alpha_1 dNPL_i^{t-2} + \alpha_2 dNPL_i^{t-1} + \alpha_3 dNPL_i^t \\ &+ \alpha_4 dNPL_i^{t+1} + \alpha_5 loang_i^t + \theta_i + \delta_t + \epsilon_i^t \end{aligned} \tag{9}$$

The second measure of opacity, *Opacity2*, is then estimated as the natural log of absolute residuals from Eq. (9).

3.2.4. Control Variables

We employ standard control variables identified in the literature. We use *Size*, which is the natural logarithm of total assets; profitability proxied by *ROA*, which is net income scaled by total assets; leverage, *LEV*, which is the ratio of total equity to total assets; market-to-book ratio, *MB*, which is the ratio of market capitalization to book equity; non-performing loans ratio, calculated as total non-performing loans scaled by total loans; net interest margin, *NIM*; industry competition proxied by the Herfindahl-Hirschman Index, *HHI*; prior three-month cumulative returns, *RET*; and detrended equity turnover, *DTURN*. *DTURN* controls for heterogeneity in investor trading patterns, which is argued to be a significant predictor of future crash risk (Hong and Stein, 2003). We have appended the definition of all variables employed in this study in Appendix 1.

3.3. Empirical strategy

Our primary specification is a logit regression model as in Eq.(10):

$$DV_i^t = \beta_0 + \beta_1 \Delta PD_i^{t-1} + \sum_k \beta_k X_i^{t-1} + \epsilon_i^t \tag{10}$$

Here, DV_i^t is our dependent variable which takes on the measures of crash risk for bank *i* in month *t*. In our primary (logit) specification, this variable is *c309*. X_i^{t-1} is the vector of

[7] This data is retrieved from the website of the Federal Reserve Bank of St. Louis: https://fred.stlouisfed.org/

control variables as outlined in Section 3.2.4. The main explanatory variable is ΔPD_i^{t-1}, which is the 3-month change in the probability of default of bank i in month *t-1.* All explanatory and control variables are lagged by one month. Eq.(10) is estimated using bank and year dummies to control for potential unobservable effects specific to the bank or year level. We check the robustness of our baseline specification using continuous measures of crash risk, wherein we estimate Eq.(8) using bank and year-fixed effects. To avoid the possibility of results being confounded by outliers, all continuous variables are winsorized at 1 and 99 percentiles. We also standardize all continuous variables to a mean of zero and a standard deviation of one. Further, all our specifications estimate standard errors clustered at the bank-level.

4. Preliminary analysis

The preliminary analysis is conducted in two ways. First, we examine the summary statistics of the distress risk, crash risk, and other control variables, as well as the association between the variables. Second, we explore the movement of distress risk and crash risk of stock prices.

4.1. Summary statistics

Table 1 presents summary statistics based on our panel of 1,20,045 bank-months, with our crash measures in Panel A and other variables in Panel B. The primary crash measure, *c309*, has a mean of 0.023 and a standard deviation of 0.151. The estimates of *c320* are similar. These statistics are comparable with prior crash risk literature using monthly panels (Andreou et al. 2021). The statistical properties of continuous crash risk measures, namely *NCSKEW* and *DUVOL*, are also comparable with the literature. The primary explanatory variable, *ΔPD,* has a mean of -0.003 and a standard deviation of 0.229. The mean bank size (natural logarithm of total assets) is 7.568. With respect to banking indicators, we note that the mean net interest margin (*NIM*) is 3.66, whereas the mean non-performing loans to total loans ratio is 1.41%.

Table 2 presents Pearson correlation measures. In Panel A, we observe that the binary crash risk measures (c309 and c320) are highly correlated with a coefficient of 0.92. NCSKEW and DUVOL also have a significant positive correlation of 0.88. The correlation between the binary and continuous measures is significantly positive, albeit lower. This is in line with the arguments of Andreou et al. (2021), wherein the binary measures only capture the actual incidence of tail risk, whereas continuous measures reflect the skewness of the return distribution (Chen et al., 2001; Jin and Myers, 2006). More importantly, we observe a positive and significant correlation (0.073) between *c309* and *ΔPD,* which is the first indication of a positive association between a build-up of bank distress and future crash likelihood. We also find that crash is positively correlated with DTURN and negatively correlated with RET. This is again consistent with prior literature (Andreou et al. 2021; Chang et al. 2017; Kim et al. 2016;). Profitability, proxied by return-on-assets (ROA), is found to be negatively correlated with crash risk. In contrast, opacity, a proxy for possible earnings manipulation by banks, is positively correlated with crash risk. These are consistent with our expectations and with the general crash risk literature (Andreou et al., 2017; Dewally and Shao, 2012). We also find a significant positive correlation between our two measures of *opacity* (0.426).

4.2. Exploratory analysis

Prior to a formal investigation of a distress-crash risk association, we graphically examine the behavior of bank distress around a month in which the bank experienced a crash (crash month). For this, we first select banks that have experienced at least one crash over the entire period under study. We then plot the average probability of default (distress) over the 12 months leading up to and after the crash month (Fig.2). Our period of study also covers two major periods of systemic or market-wide distress: the global financial crisis (GFC), which encompassed the year 2008; and the Covid-19 pandemic in 2020. To ensure our results are not being driven by possibly higher distress during these periods, Fig.2 also shows the plot excluding 2008 and 2020 as well. While default probability is observed to be relatively low and stable about 8 months prior to the crash, it begins to increase about 6 months pre-crash. The trends are similar even when we exclude the years 2008 and 2020. We find that default probability remains high as long as six months post-crash, after which it begins to decline. This may indicate that, while stock price crashes may occur because of a sudden revelation of unfavourable news, there is a considerable period before the actual crash during which the leakage of this news becomes gradually evident, which results in a gradual increase in distress risk. This phenomenon also indicates support for the "bad news hoarding" mechanism (Kim et al. 2011), which attributes stock price crashes primarily to managers' opportunistic withholding of unfavorable news and its sudden revelation to the market.

Table 1: Summary Statistics

Variable	Obs.	Mean	SD	Min	Q1	Median	Q3	Max
				Panel A: Crash Risk Measures				
c309	120045	0.023	0.151	0.000	0.000	0.000	0.000	1
c320	120045	0.020	0.140	0.000	0.000	0.000	0.000	1
NCSKEW	120045	0.066	0.736	-2.163	-0.339	0.046	0.451	2.485
DUVOL	120045	0.049	0.491	-1.194	-0.267	0.042	0.361	1.416
				Panel B: Other Variables				
ΔPD	118292	-0.003	0.229	-0.778	-0.037	0.000	0.035	0.743
ΔPR	118292	-0.003	0.260	-0.829	-0.066	0.000	0.064	0.771
Size	120045	7.568	1.495	4.979	6.500	7.255	8.388	12.489
ROA	120045	0.001	0.002	-0.012	0.001	0.002	0.003	0.005
LEV	120045	0.100	0.028	0.040	0.081	0.096	0.114	0.205
MB	120045	1.370	0.651	0.208	0.934	1.243	1.694	3.650
RET	120045	0.006	0.152	-0.556	-0.058	0.010	0.082	0.438
DTURN	120045	0.0001	0.001	-0.004	-0.0005	-0.0001	0.0003	0.0075
HHI	120045	0.198	1.339	0.000	0.0001	0.0003	0.0028	12.085
MDZ	119971	68.299	43.235	9.688	40.263	59.802	84.711	265.113
NIM	120045	3.660	0.729	1.960	3.200	3.600	4.050	5.950
NPL	115372	1.418	1.758	0.021	0.391	0.795	1.670	10.100
LLP	115282	0.001	0.002	-0.001	0.000	0.000	0.001	0.015
Opacity1	110375	-7.391	1.250	-11.327	-8.068	-7.243	-6.588	-4.491
Opacity2	109569	-7.552	1.299	-11.568	-8.270	-7.448	-6.695	-4.624

Note: This table presents summary statistics of all major variables. All variables pertain to bank-month panel over the period 2000-2022. Variables are winsorized at 1% and 99% of their distribution. Variable descriptions are available in Appendix 1.

Table 2: Pearson correlation matrix

Panel (A): correlation between crash risks

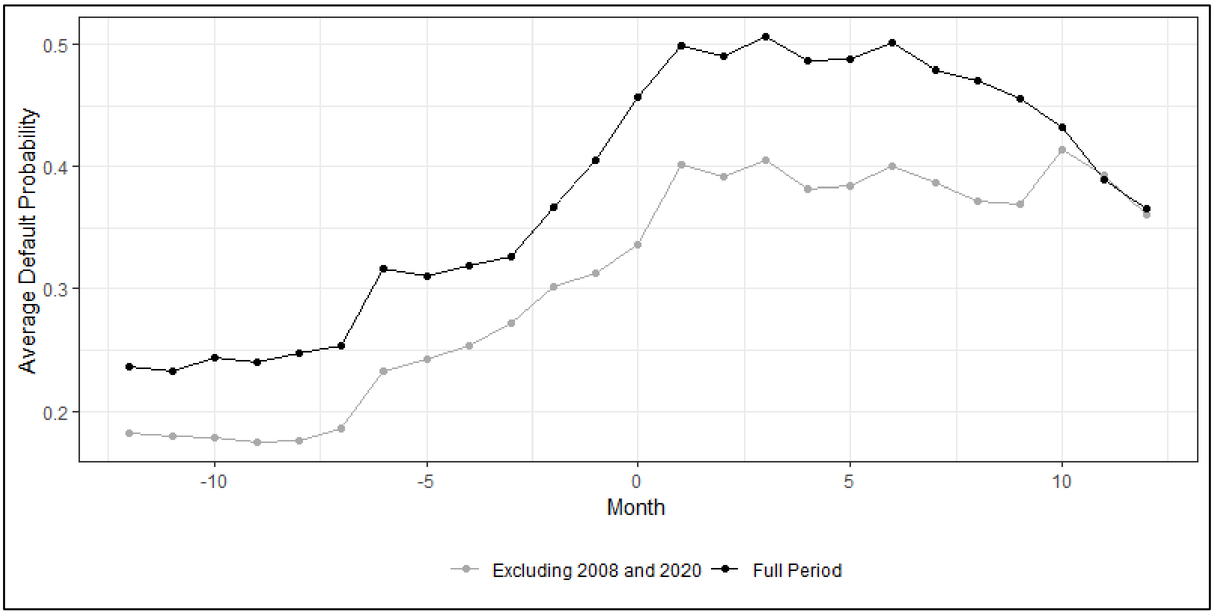
	$c309_t$	*$c320_t$*	*NCSKEW*	*DUVOL*
$c309_t$	1			
$c320_t$	0.926**	1		
NCSKEW	0.189**	0.191**	1	
DUVOL	0.159**	0.157**	0.888**	1

Panel (B): correlation between crash risk and other bank-specific variables

	$c3.09_t$	*ΔPD*	*Size*	*MB*	*Lev*	*ROA*	*Ret*	*DTURN*	*Opacity1*	*Opacity2*
$c309_t$	1									
ΔPD	0.073**	1								
Size	-0.005	0.011**	1							
M/B	-0.026**	-0.008**	0.840**	1						
Lev	-0.016**	0.014**	0.515**	-0.070**	1					
ROA	-0.067**	-0.029**	0.101**	0.168**	0.149**	1				
Ret	-0.174**	-0.344**	0.001	0.124**	0.045**	0.255**	1			
DTURN	0.039**	0.010**	0.004	0.011**	-0.008**	-0.016**	0.012**	1		
Opacity 1	0.032**	0.015**	0.082**	-0.086**	-0.022**	-0.159**	-0.113**	0.0011	1	
Opacity 2	0.0346**	0.0042	-0.025**	-0.098**	-0.026**	-0.186**	-0.133**	0.0055	0.426**	1

Note: This table presents Pearson correlation coefficients between all major variables. *, **, and *** indicate statistical significance at 10%, 5%, and 1% confidence respectively.

Fig. 2: Behaviour of probability of default surrounding a crash



Note: This figure shows the evolution of the probability of default of 'crash' banks in the 12 months prior to and post a crash month. A crash month is defined as a month during which the bank has experienced at least one bank-specific weekly return, which was 3.09 standard deviations below the mean bank-specific weekly return over the estimation window. Crash month is indicated by month 0. Crash banks constitute banks that have faced at least one crash over the entire sample period. The plot in black is for the entire sample period (2000-2022), whereas the plot in grey excludes two years: the GFC of 2008 and the Covid-19 pandemic of 2020

5. Results and discussion

We now formally examine the association between an increase in default probability and future crash risk. Estimates from the baseline logit regression, as in Eq. (8), are presented in Table 3. The dependent variable is *c309*, the main binary crash risk measure. Specification 1 does not include any control variables. Consistent with our expectations and previous analysis, we find a significantly positive association between a three-month change in probability of default (*ΔPD)* and future crash risk. This association continues to hold when we gradually add control variables in specifications 2, 3, and 4. All four specifications include bank and year dummies to control for omitted variable bias arising out of potential unobserved factors affecting crash risk. In specification 5, we control for industry-fixed effects instead of bank-fixed effects, where industries are classified according to the Fama-French 48 industries. We continue to find a significant positive effect of a three-month increase in the probability of default and future crash risk. Importantly, we find that the coefficient on *ΔPD* does not vary much across specifications 2-5, which highlights the significance of increasing distress as a precursor to stock price crashes.

Table 3: Baseline Regressions: Probability of default and stock price crash risk

Variables	(1) c309$_t$	(2) c309$_t$	(3) c309$_t$	Marginal effects	(4) c309$_t$	(5) c309$_t$
ΔPD_{t-1}	0.203***	0.116***	0.107***	0.0029***	0.119***	0.109***
	(0.018)	(0.019)	(0.019)		(0.019)	(0.019)
$Size_{t-1}$		0.065*	-0.003	-0.00008	0.0011	0.085***
		(0.033)	(0.043)		(0.045)	(0.030)
Lev_{t-1}		0.013	0.018	0.0005	0.012	0.024
		(0.021)	(0.021)		(0.022)	(0.021)
M/B_{t-1}		0.275***	0.284***	0.0068***	0.289***	0.249***
		(0.029)	(0.029)		(0.031)	(0.028)
ROA_{t-1}		-0.051**	-0.049**	-0.0012**	-0.0497**	-0.043**
		(0.021)	(0.021)		(0.023)	(0.021)
Ret_{t-1}		-0.256***	-0.246***	-0.0059***	-0.247***	-0.227***
		(0.025)	(0.024)		(0.025)	(0.024)
$DTURN_{t-1}$		0.100***	0.0900***	0.0022***	0.0939***	0.086***
		(0.016)	(0.017)		(0.017)	(0.016)
HHI_{t-1}			0.0880***	0.0021***	0.0859***	0.057**
			(0.026)		(0.027)	(0.023)
MDZ_{t-1}				0.00002***	-0.00261**	-0.003***
					(0.001)	(0.001)
NIM_{t-1}					0.0320	0.032
					(0.024)	(0.023)
NPL_{t-1}					0.034	0.061**
					(0.026)	(0.025)
Bank FE	YES	YES	YES		YES	
Industry FE						YES
Year FE	YES	YES	YES		YES	YES
Constant	-3.820***	-3.690***	-3.716***		-3.779***	-3.761***
	(0.202)	(0.227)	(0.257)		(0.273)	(0.218)
Observations	105,622	104,873	104,309		99,900	109,301

Note: This table presents coefficient estimates from logit regressions of the form specified in Eq.10. The dependent variable in all specifications is *c309*, the binary measure of crash risk, which takes a value of 1 if in that month the bank has experienced a bank-specific return which is lower than 3.09 standard deviations from the mean bank-specific return over the estimation window. All continuous variables have been winsorized at 1 and 99 percentiles and standardized to have a mean of zero and a standard deviation of one. Specifications 1-4 include bank and year dummies to control for unobserved heterogeneity, whereas specification 5 includes industry dummies constructed according to the Fama-French 48 industry classification. Values in parentheses are standard errors clustered at bank level. *, **, and *** indicate statistical significance at 10%, 5%, and 1% confidence level respectively.

The coefficients on the control variable are broadly consistent with prior literature. Consistent with Dewally and Shao (2013), larger banks have a higher crash risk, as indicated by a significant positive coefficient on s*ize* (specifications 2 and 5). Banks with a higher market-to-book ratio are also more vulnerable to future crashes, as evidenced by a positive and significant coefficient on *MB*. This also conforms to existing literature on bank stock price crash risk (PC Andreou et al. 2017). In line with the previous studies (Andreou et al., 2021; Kim et al., 2021;

and Kim and Zhan, 2016), we find that crash risk is lower for banks with higher prior-period cumulative return (*RET)* and higher for banks with more investor belief heterogeneity (proxied by higher detrended equity turnover, *DTURN*). We do not find any evidence of crash risk being affected by net interest margin (NIM), and only a weak positive association with non-performing loans ratio (*NPLRATIO)* (specification 5). Overall, we continue to find a positive and significant association between change in default probability and future crash risk across multiple specifications and multiple controls.

6. Robustness tests

We conduct a series of robustness analyses to cement our baseline results. These tests include employing alternate default probability and crash risk measures, analysing DD-component variables, and tests for reverse causality.

6.1. Alternative measures of default and crash risks

The distance-to-default has been widely employed in the literature on firm distress, both for non-financial firms (Andreou et al., 2021; Bharath and Shumway, 2008; Islam et al., 2021), as well as for banks and financial institutions (Harada et al., 2013; Nagel and Purnanandam, 2020). However, an incidence of default, or a build-up to it, may entail widely different ramifications for the economy if the defaulting firm is a banking institution. Essentially, bank defaults differ from corporate defaults along two fronts. First, banking institutions have a different liability structure compared to non-financial firms. Banks are funded to a large extent by both short- and long-term external deposits, a major chunk of which comes from retail depositors (public money). Banks, thus, have a much higher leverage than non-financial firms, which is attributable to their business model. The naïve DD model, however, does not take this into account and, as such, would always predict a much higher probability of default. Second, the costs associated with a bank default are substantially larger than a corporate default (Hoelscher and Quintyn, 2003). Being at the core of the economy, banks are heavily intertwined with the global financial system. An incidence of bank default not only affects all depositing and borrowing entities directly connected to it but has the potential to incite widespread panic and bank-runs at other financial institutions, resulting in a systemic financial crisis. Regulators and governments thus have an incentive to bail out even idiosyncratic bank failures to avoid the possibility of a larger systemic event. The costs of bailout by the government or other financial institutions may be substantial. Government bailouts in the United States in the aftermath of the GFC are pegged at USD500 Billion, or about 3.5% of US GDP in 2009 (Lucas, 2019). Hence, even when a bank is not on the verge of default, any incidence of financial weakness attracts intervention by regulators in the form of supervision, corrective action, and stringent capital requirements. As the Basel Committee on Banking Supervision proposed, such actions may be triggered when the bank falls short of mandatory capital requirements. Any bank failing to meet the required capital requirements is placed under a "prompt corrective action" framework (PCA), which imposes restrictions on banking activities and regulatory supervision to recapitalize the bank. The naïve *DTD* methodology does not take into account the intricacies associated with the functioning of banking institutions. These arguments are put forward by Chan-Lau and Sy (2006), who then propose an alternate measure of default risk for banks, the distance-to-capital (*DC*). They propose that banks' default threshold (total liabilities) must be augmented with the capital requirement under the PCA framework in place. Specifically, they propose that total liabilities for banks must be multiplied by an additional factor λ:

$$\lambda = \frac{1}{(1 - PCAR_i)} \tag{11}$$

Hence, as an alternate measure of distance to default, we estimate the distance to capital and the corresponding probability of risk (*PR*)[8] for the banks in our study. We re-estimate our baseline model using the corresponding three-month change in probability of risk (*ΔPR).* We also consider three alternate crash risk measures: another binary indicator, *c320*, and two continuous measures *NCSKEW* and *DUVOL.* Column 1, Panel A, in Table 4 shows the baseline regression estimates with *ΔPD* as the explanatory variable, whereas columns 2 and 3 have *ΔPR* as the explanatory variable. Column 2 employs bank and year fixed effects, whereas we use industry fixed effects in column 3. All three specifications have c309 as the crash risk measure and include all baseline control variables. We continue to find support for a positive *ΔPR-crash* association, as evidenced by a statistically significant positive coefficient in both specifications.

The association continues to hold when we use c320 as the measure of crash risk in columns 4 and 5, Panel A. The results are reiterated when we use the continuous measures of crash risk in Panel B. These findings indicate that a positive default risk-crash risk association is robust to alternate measures of crash as well as default.

[8] Probability of risk (PR) is estimated in a manner similar to probability of default (PD). While PD is estimated from distance to default (DD) as *PD=N(-DD),* PR is estimated from distance to capital (DC) as *PR=N(-DC).*

Table 4: Robustness: Alternative Measures of Default probability and crash risk

Note: This table presents coefficient estimates using alternative measures of default probability and stock price crash risk. Panel A shows the results of logit regressions of the form specified in Eq. 10, with year-fixed effects and bank/industry fixed effects. *ΔPR* is the alternative measure of default probability estimated using the distance-to-capital proposed by Chan-Lau and Sy (2007). The dependent variable in Panel A is the alternative crash risk measure, *C320*, which takes a value of 1 in the month when the bank has a bank-specific weekly return that is 3.20 standard deviations lower than the mean bank-specific weekly return over the estimation window. Panel B shows the results of OLS regressions with bank and year fixed effects, where the dependent variables are two alternative continuous crash risk measures, *NCSKEW* and *DUVOL*. All continuous variables have been winsorized at 1 and 99 percentiles and standardized to have a mean of zero and a standard deviation of one. Values in parentheses are standard errors clustered at bank level. *, **, and *** indicate statistical significance at 10%, 5%, and 1% confidence level respectively.

Panel (A): Distance to capital and crash risk at 3.20					
	$c309_t$	$c309_t$	$c309_t$	$c320_t$	$c320_t$
ΔPD_{t-1}	0.109*** (0.020)			0.137*** (0.020)	
ΔPR_{t-1}		0.084*** (0.021)	0.075*** (0.020)		0.099*** (0.0210)
Controls	YES	YES	YES	YES	YES
Bank FE	YES	YES		YES	YES
Industry FE			YES		
Year FE	YES	YES	YES	YES	YES
Obs.	109,301	99,900	109,301	97,874	97,874
Panel (B): *NCSKEW* and *DUVOL*					
	NCSKEW			*DUVOL*	
ΔPD_{t-1}	0.009*** (0.002)			0.008*** (0.001)	
ΔPR_{t-1}		0.009*** (0.002)			0.007*** (0.001)
Controls	YES	YES		YES	YES
Bank FE	YES	YES		YES	YES
Year FE	YES	YES		YES	YES
R^2	0.630	0.630		0.646	0.646
Obs.	109,314	109,314		109,317	109,317

6.2. Components of Distance to Default

To ensure that the association between *ΔPD* and crash risk is not driven specifically by any of its components but rather the default probability as a whole, we re-estimate the baseline models after including each of the three components of *DTD* as additional controls: (a) 3-month cumulative returns (AR_i^{t-1}), (b) ratio of bank value to total liabilities (VD_{t-1}), and (c) asset volatility (σ_{AS}). The regression estimates are presented in Table 5. Column 1 is our baseline specification without DD components, and we gradually add each component one by one in specifications 2-4 until specification 5 which includes all three components. We continue to find a strong positive coefficient on *ΔPD,* which shows that it is the default probability as a whole that continues to reveal information about crash risk, and this association is unlikely to be driven by any of the component variables.

Table 5: Analyses of Distance-to-default components

	(1)	(2)	(3)	(4)	(5)
	$c309_t$	$c309_t$	$c309_t$	$c309_t$	$c309_t$
ΔPD_{t-1}	0.109***	0.114***	0.115***	0.104***	0.105***
	(0.020)	(0.0196)	(0.0195)	(0.0193)	(0.0197)
AR_{t-2}		0.236***			0.0967
		(0.0713)			(0.0849)
VD_{t-1}			0.00002		0.00003
			(-0.00002)		-0.00002
$\sigma_{AS,t-1}$				-3.983***	-3.748***
				(0.722)	(0.760)
Controls	YES	YES	YES	YES	YES
Bank FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES
Obs.	109,301	98,297	98193	99,900	96,647

Note: This table presents coefficient estimates from our baseline logit regressions augmented with *DTD* components as additional control variables. AR_{t-2} is the cumulative equity return of a bank over the 12-month period, two months prior to a crash. VD_{t-1} is the ratio of the bank's total value to its liabilities' book value. $\sigma_{AS,t-1}$ is the volatility of the bank's assets. All specifications also include the baseline control variables and bank and year dummies to control for any omitted variable bias arising due to unobserved heterogeneity. Values in parentheses are standard errors clustered at bank level. *, **, and *** indicate statistical significance at 10%, 5%, and 1% confidence level respectively.

There may be a possibility of our results being affected by unobserved factors, which may introduce endogeneity in our model. To alleviate such concerns, our baseline and robustness models in Tables 3 and 4 include year-fixed effects as well as bank and/or industry-fixed effects. We also include all standard controls as employed in the stock price crash risk literature. Our results are consistent across the inclusion of fixed effects as well as the employment of multiple crash risk and default risk measures.

6.3. Issue of reverse causality

Our baseline regressions employ one-month lagged *ΔPD* and lagged control variables in all specifications. A lead-lagged relationship may help alleviate concerns of possible reverse causality. Moreover, Fig.2 also evidences a dramatic increase in default probability in the six months just prior to a stock price crash. To strengthen our results further, we follow Andreou et al. (2021) and conduct three additional tests to reduce the possibility of our results being affected by reverse causality.

First, we re-estimate the baseline regression using all four crash measures *(c309, c320, NCSKEW, and DUVOL*) after using three lags of crash risk as additional controls. The coefficient estimates, presented in Table 6, Panel A, show that 1-month lagged *ΔPD* continues to have a statistically significant positive association with crash risk, even after the possible effects of prior period crash risk have been controlled for. Moreover, we find that the magnitude of coefficients remains qualitatively similar to our baseline models.

Next, to examine whether the crash risk results in an increase in the probability of default, we estimate OLS regressions, keeping the probability of default (*ΔPD)* as the dependent variable and lagged crash risk as the main explanatory variable. We use up to five lags of crash risk as the explanatory variable. Panel B of Table 6 presents the coefficient estimates. We find that only one-period lagged crash risk has a significant positive effect on *ΔPD.* This is to be expected for two reasons. First, a stock price crash depresses equity market capitalization, which in turn reduces firm value. This may temporarily increase default probability. Second, a stock price crash may induce a heavy sell-off by institutional investors, which in turn induces panic selling among other investors (Chang et al. 2017). The ultimate effect of such a sell-off is decreased market equity, further aggravating default probability (Andreou et al. 2021). However, we note that this effect is only temporary, as further lags (2-5) of crash risk do not appear to affect *ΔPD* significantly. These observations also echo Fig. 2, where we observe a sharp increase in default probability in the month immediately following a stock price crash. However, Fig.1 also shows that default probability begins to reduce in the following months, which is consistent with our formal regression tests.

As a final test against reverse causality and as a follow-through to the prior test, we re-run the baseline regressions after excluding data pertaining to the months immediately following a crash. The models are re-estimated after dropping observations simultaneously in the 1st, 2nd, 3rd, 4th, and 5th months after a crash. This helps add credibility to our results by weakening the possibility that an incidence of crash is driving an increase in default probability. We employ all four measures of crash risk but only report the estimates for the main binary measure (c309) for brevity. We find that a strong positive association of *ΔPD* on future crash risk continues to hold across both binary measures of crash risk (Table 6, Panel C). Coefficient estimates are qualitatively similar when we use c320 and the continuous crash measures. We conclude that our results appear robust to any endogeneity concerns arising out of reverse causality.

Table 6: Results of reverse causality

Panel A: Using three lags of crash risk

	(1)	(2)	(3)	(4)
	c309	*c320*	*NCSKEW*	*DUVOL*
ΔPD_{t-1}	0.145***	0.161***	0.007***	0.006***
	(0.020)	(0.020)	(0.001)	(0.001)
Lags of crash risk	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Bank FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Obs.	96,680	94,907	1,05,085	1,05,086
Pseudo/Adj. R2	0.094	0.101	0.632	0.648

Panel B: Effect of crash risk on default probability

	(1)	(2)	(3)	(4)	(5)
	ΔPD	*ΔPD*	*ΔPD*	*ΔPD*	*ΔPD*
$c309_{t-1}$	0.023***				
	(0.005)				
$c309_{t-2}$		0.006			
		(0.004)			
$c309_{t-3}$			0.006		
			(0.004)		
$c309_{t-4}$				-0.001	
				(0.004)	
$c309_{t-5}$					0.005
					(0.004)
Controls	Yes	Yes	Yes	Yes	Yes
Bank FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Obs.	109,827	107,700	105,602	104,435	103,444
R2	0.323	0.325	0.326	0.327	0.327

Panel C: Re-estimation of baseline regressions after excluding data in the month(s) immediately after a crash

	(1)	(2)	(3)	(4)	(5)
	c309	*c309*	*c309*	*c309*	*c309*
	1 Month	2 Months	3 Months	4 Months	5 Months
ΔPD_{t-1}	0.126***	0.138***	0.153***	0.153***	0.152***
	(0.021)	(0.022)	(0.023)	(0.023)	(0.024)
Controls	Yes	Yes	Yes	Yes	Yes
Bank FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Obs.	93,453	89,619	86,885	83,899	80,902
Pseudo/Adj. R2	0.091	0.093	0.096	0.099	0.104

Note: This table presents results from three sets of analyses on the examination of reverse causality. Panel A presents estimated from the baseline regressions where three lags of crash risk have been added as additional controls. Columns 1-2 are logit regressions, whereas columns 3-4 are estimated from OLS regressions. Panel B presents coefficient estimates from OLS regressions where the crash measure *c309* and increase in default probability *ΔPD* swap places. Panel C presents estimates from the baseline logit regressions after excluding observations, simultaneously in months 1-5 after a crash month. All specifications include bank and year-fixed effects as specified. Values in parentheses are standard errors clustered at bank level. *, **, and *** indicate statistical significance at 10%, 5%, and 1% confidence level respectively.

6.4. Propensity Score Matching

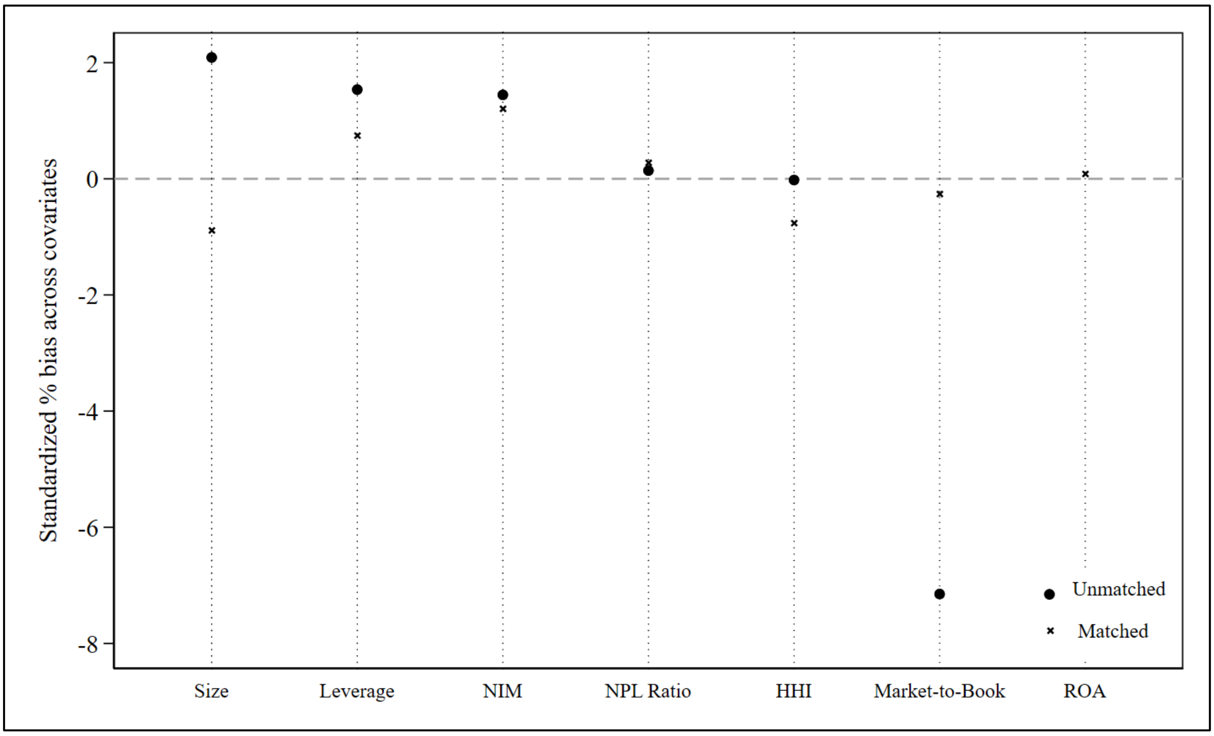
Although we included fixed effects and standard control variables in all our specifications, there still exists a possibility of an omitted-variable bias, which may arise due to possible functional form misspecification. This may be due to, for instance, a non-linear association between crash risk and one or more control variables. In that case, a significant association between **crash risk and *ΔPD*** could be because of inherent dissimilarity (with respect to one or more controls) among banks with a high ΔPD and those without. We relax the functional form assumption to alleviate such concerns by adopting the non-parametric propensity score matching method. The PSM framework allows for estimating an "average treatment effect" by matching treated groups with control (counterfactual) groups across other observables, thus effectively reducing any selection bias. The 'treated' ('control') group in our study comprises banks that had a prior-month value of *ΔPD* which was above (below) the 75th (25th) percentile of their distribution across the sample during that month. As a precursor to formal PSM analysis, we first examine whether the average treatment effect on the treated (*ATT*) is significant.[9] First, we estimate a logit model to predict the probability of a bank being treated (the 'propensity score'), conditional on the given covariates; following this, it matches treated banks to control banks that are similar in the propensity score. Finally, it gives an estimate of the *ATT* based on a comparison of the treated and control groups in the matched sample. The command estimates the *ATT* through an OLS regression of the outcome variable; thus, we estimate *ATT* using the continuous crash risk measure, *NCSKEW*. Fig. 3 shows the standardized bias across the matching covariates after psmatch2. We observe that the post-match bias is close to zero, which indicates that banks in the matched sample are adequately similar across the covariates. The *ATT* comes out to be 0.125, with a standard error of 0.008 (*t-statistic* = 15.07), which is the first indication that functional form misspecification is not an issue in our formal results.

Given that our data is a bank-month sample, it becomes imperative to ensure matching treated banks with control banks within the same month or year to ensure that the matching accounts for any time-specific heterogeneity. To this end, we conduct an additional PSM analysis. We follow three steps. First, we identify treatment and control groups. As our treatment variable, *ΔPD,* is a continuous variable, for each year-month, we split the value of *ΔPD* into four quartiles. The treatment variable takes a value of 1 for banks that have a value

[9] We use the psmatch2 program in Stata.

of *ΔPD* in the preceding month equal to or above the 75th percentile (high *ΔPD).* The control group (treatment variable = 0) are banks whose *ΔPD* in the preceding month is equal to or below the 25th percentile. We base our treatment variable on one-month lagged *ΔPD* because we are examining the effect of one-month lagged *ΔPD* on crash risk. Next, we estimate a logit regression for the treatment variable on the main fundamental observable control variables, all lagged by a month: Size, ROA, Leverage, Market-to-book, HHI, net-interest-margin, and non-performing loans ratio. Logit regression aims to extract the likelihood of a bank being treated, which is the propensity score. Having extracted the propensity score, we then match each treated bank with a control bank within the same year-month, such that they have similar propensity scores, which essentially means a similar likelihood of being treated. We employ a caliper distance of 0.05 for matching, in line with Kang et al. (2019). Finally, we re-estimate our baseline regression on the matched sample. Table 7 presents the results of PSM. To ensure that banks in the matched sample are similar across the controls, we compare coefficient estimates from a logit regression on the full sample with that on the matched sample (Panel A, Table 7). We find that, while there was significant heterogeneity across controls in the full sample, there are no significant differences in the matched sample, which lends credibility to our matching process. We present the coefficient estimates from regressions on the matched samples in Panel B. The positive coefficients on *ΔPD* across all four crash risk measures support our baseline results and weaken any concerns of biased estimates due to functional-form misspecification.

Fig. 3: Bias Across Covariates: Unmatched vs Matched Sample



Note: This figure depicts the standardized percentage bias in covariates for the unmatched sample and the matched sample, generated using the *-psmatch2-* command in Stata.

6.5. Examination of bank opacity around default risk

The general crash risk literature posits managerial tendencies to withhold unfavourable news as a major driver of stock price crash risk (Andreou et al. 2017; Kim et al. 2011). The argument stems from the classical agency theory (Jensen and Meckling, 1976), which posited that separation of ownership and control in publicly traded firms results in a misalignment in the incentives of managers and shareholders/investors. While shareholders/investors would focus on maximization of firm value, managers may not prioritize the same as they may be driven by other incentives such as maximizing personal wealth and career concerns. Thus, managers may be inclined to consciously withhold bad news about a firm's financial health by managing earnings in a favorable way. This exacerbates information asymmetry among managers and external stakeholders by reducing the transparency of financial reporting and ultimately results in a stock price crash when the withheld information reaches a maximum. Consistent with these arguments, Andreou et al. (2021) document that, for non-financial firms in the US, firm opacity is significantly higher after a firm faces an increase in default risk, which in turn results in higher stock price crash risk.

Table 7: Propensity Score Matching

	Panel A: Matching Diagnostics			**Panel B: Regression on Matched Sample**			
	(1)	*(2)*		*(3)*	*(4)*	*(5)*	*(6)*
	Full Sample	*Matched Sample*		*c309*	*c320*	*NCSKEW*	*DUVOL*
				0.108***	0.114***	0.009***	0.008***
Size	0.163*** (0.040)	-0.055 (0.059)	*ΔPD(t-1)*	(0.036)	(0.038)	(0.002)	(0.001)
ROA	-56.73*** (7.588)	-5.727 (11.841)					
LEV	-0.981* (0.704)	-0.936 (1.141)	*Controls*	Yes	Yes	Yes	Yes
MB	-0.433*** (0.038)	0.076 (0.061)	*Bank FE*	Yes	Yes	Yes	Yes
HHI	0.016 (0.037)	-0.045 (0.070)	*Year Dummies*	Yes	Yes	Yes	Yes
NIM	0.184*** (0.028)	-0.018 (0.044)	*Observations*	17,090	15,745	27,715	27,715
NPL	-0.041*** (0.011)	0.000 (0.016)	*Pseudo/Adj. R2*	0.118	0.130	0.629	0.639
Bank FE	Yes	Yes					
Year FE	Yes	Yes					
Obs.	55,434	28,001					

Note: This table presents the results from propensity score matching analyses. Panel A shows diagnostics before and after propensity score matching, where the dependent variable in columns 1-2 is the treatment dummy, which takes a value of 1 for bank months if, in the prior month, the bank's *ΔPD* was above the 75^{th} percentile of all banks in that month., and 0 otherwise. Column 1 is for the full sample, whereas column 2 is only for the matched sample. In Panel B, the coefficient estimates are from logit (columns 3-4) and OLS (columns 5-6) regressions of crash risk on lagged *ΔPD* and all baseline controls for the matched sample. All variables are winsorized at 1% and 99% of the distribution. Values in parentheses are standard errors clustered at bank level. *, **, and *** indicate statistical significance at 10%, 5%, and 1% respectively.

Concerns about earnings management are especially important for banking firms because, unlike non-financial firms, a bank's assets (loan portfolio) are difficult to value by external investors, making them inherently opaque (Campbell and Kracaw, 1980; Morgan, 2002). More recently, Blau et al. (2017) find that the efficiency of stock prices is significantly lower for banks than non-banks, which they attribute to higher opacity of banks. Hence, we next examine if and how bank opacity may be a channel through which an increase in default probability affects stock price crash risk. We first follow a procedure similar to Andreou et al. (2021) and conduct the analyses in two stages. In the first stage, we predict the three-month change in bank opacity from the following regression:

$$\Delta Opacity_i^t = \alpha_0 + \alpha_1 \Delta PD_i^{t-1} + \sum_k \alpha_k X_i^{t-1} + \theta_i + \delta_t + \epsilon_i^t \qquad (12)$$

$\Delta Opacity$ is the three-month change in bank opacity, where we employ the second measure of opacity, as proposed by Zheng (2020), as the measure of bank opacity. This model considers the *absolute value* of residuals from a model of loan loss provisions (*LLP*) as a measure of opaqueness, or discretion exercised by managers, in reporting *LLP*. Table 7, Panel A shows the coefficient estimates from the first stage regression. We find that banks that face an increase in default risk are associated with a higher value of opacity in the next period, which is indicative of a higher tendency to manage reported *LLPs* in the face of mounting distress. To examine whether the higher probability of crash risk is driven, in part, by higher opacity, we extract the predicted values of $\Delta Opacity$ from the first stage regression, and, in the second stage, we model the relationship between opacity conditional on the increase in default risk, and future stock price crash risk, via a logit model with bank and year fixed effects. Contrary to Andreou et al. (2021), however, we find a negative association between opacity conditional on increasing default risk and future stock price crash risk (Table 7, Panel B). This appears counterintuitive regarding the broader crash risk literature, as it suggests that higher opacity, and thus, higher earnings management, *reduces* crash risk.

In our attempt to investigate this apparent aberration from the extant literature, we delve deeper into the specific literature on discretionary loan loss provisioning in the banking sector. Our measure of opacity in the previous section was based on *absolute* residuals from a model of *LLPs*. These residuals are the discretionary component of *LLPs* (*DLLPs*), and hence, have been proposed as a measure of opacity. The rationale behind absolute and not raw values of residuals is that both higher and lower discretionary *LLPs* (*DLLPs*) are indicative of earnings management and, thus, of opacity (Jiang et al., 2016; Zheng, 2020). However, a considerable body of literature argues that the level of *DLLPs* is affected by different managerial motivations. Bank management may, for instance, show a higher or lower *DLLP* to smoothen earnings over reporting periods because earnings volatility may draw a negative reaction from the stock market (Ma, 1988). Evidence of bank earnings smoothing through *DLLPs* is provided by Bushman and Williams (2012), Kanagaretnam et al. (2003), Laeven and Majnoni (2003), and Lobo and Yang (2001). Another motivation for higher or lower *DLLPs* is signaling the bank's earnings prospects to the stock market (Beaver et al., 1989; Lobo and Yang, 2001; Morris et al., 2016). Although one may expect a sudden increase in *LLPs* to reflect a deteriorating state of the bank's assets, extant literature provides evidence to the contrary. The literature argues that a higher amount *of DLLPs* signals to the market that the bank is able to build up a buffer of higher *LLPs* and is, thus, financially healthy. In addition, it is argued that higher *DLLPs* are associated with more conservative banking practices and more timely loan loss recognition and act as a buffer for the bank during periods of weaker financial health. Consistent with these arguments, empirical evidence indicates that the stock market reacts favourably to higher discretionary provisions (Beaver and Engel, 1996; Kanagaretnam et al., 2005; Wahlen, 1994). More recently, Hegde and Kozlowski (2021) find that, during weaker economic states, a sudden increase in *DLLPs* draws a negative stock market reaction because investors become wary of higher *LLPs* in the face of mounting default concerns. During good economic states, on the other hand, higher *DLLPs* draw a positive market reaction. Banks with more conservative *LLPs* (Andreou et al. 2017) and banks with more timely loss recognition (Jung et al. 2019) are also found to have lower stock price crash risk. These studies indicate that the valuation implications of positive *DLLPs* are substantially different from those of negative *DLLPs.* We thus conclude that, for banks specifically, a measure of opacity that considers the absolute value of residuals from a model of *LLPs* may not accurately reveal the association between opacity and crash risk because it does not differentiate between positive and negative *DLLPs*. Thus, we re-estimate our analyses by considering the raw residuals from *LLP* models outlined in Jiang et al. (2016) and Zheng (2020) as our new measure of opacity[10]. This is also consistent with Hegde and Kozlowski (2021) and Morris et al. (2016).

[10] The residual is the difference between the reported and the predicted *LLP*. A positive residual, thus, indicates that the bank reported higher than expected *LLPs*, and conversely for negative residuals.

Table 8: Analyses of bank opacity

	(1)	**(2)**
	$\Delta Opacity1_t$	$\Delta Opacity2_t$
Panel (A): Effects of changes in distress risk on opacity		
ΔPD_{t-1}	0.000	0.012***
	(0.005)	(0.004)
Controls	YES	YES
Bank FE	YES	YES
Year FE	YES	YES
Obs.	104,477	103,817
R2	0.095	0.172

Panel (B): $\Delta\widehat{Opacity}_t$ on Crash risk	$c309_t$
$\Delta\widehat{\text{Opacity}}_{t-1}$	-1.165***
	(0.156)
Controls	YES
Bank FE	YES
Year FE	YES
Obs.	97,449

Notes: This table presents coefficient estimates from analyses of the moderating effect of bank opacity on the association between ΔPD and crash risk. Panel A is the first stage regression of the two baseline measures of opacity on lagged ΔPD. Opacity1 corresponds to opacity estimated from the model proposed by Jiang et al. (2016), and Opacity2 corresponds to the model proposed by Zheng (2020). *ΔOpacity* is the three-month increase in opacity for a bank. Models 1 and 2 in Panel A correspond to OLS regressions with bank and year-fixed effects. Panel B presents estimates from the second stage regression of stock price crash risk on predicted estimates of opacity obtained from the first stage. The specification in Panel B is a logit model with bank and year dummies to control for unobserved heterogeneity. Values in parentheses are standard errors clustered at bank level. *, **, and *** indicate statistical significance at 10%, 5%, and 1% respectively.

Table 9: Analyses using alternative measures of opacity

Panel A: First Stage Regressions			**Panel B: Second Stage Regressions**		
	(1)	(2)		(1)	(2)
	$Opacity1_t$	$Opacity2_t$		$c309_t$	$c309_t$
ΔPD_{t-1}	-0.0072*	-0.0248***	$Opacity1_{t-1}$	-0.5420***	
	(0.0038)	(0.0043)		(0.0842)	
			$Opacity2_{t-1}$		-0.492***
					(0.0868)
Controls	Yes	Yes	Controls	Yes	Yes
Bank FE	Yes	Yes	Bank FE	Yes	Yes
Year FE	Yes	Yes	Year FE	Yes	Yes
Obs.	102,261	101,647	Obs.	92,657	92,196
R-squared	0.288	0.154	Pseudo R-squared	0.092	0.092

Notes: This table presents the results from alternative opacity measures. *Opacity1 and Opacity2* are, respectively, the raw residuals from a model of loan loss provisions as proposed by Jiang et al. (2016) and Zheng (2020). Models 1 and 2 in Panel A correspond to OLS regressions with bank and year-fixed effects. Panel B presents estimates from the second stage regression of stock price crash risk *on predicted estimates* of opacity obtained from the first stage. The specification in Panel B is a logit model with bank and year dummies to control for unobserved heterogeneity. Values in parentheses are standard errors clustered at bank level. *, **, and *** indicate statistical significance at 10%, 5%, and 1% respectively.

To examine whether an increase in default probability affects bank's opacity, we regress the opacity measures on one-month lagged ΔPD. The coefficient estimates, presented in Table 9 (Panel A), reveal that the two are negatively associated. In the face of increasing default risk, we conjecture that banks reduce their discretionary $LLPs$, resulting in lower-than-expected reported LLP. This may indicate a managerial tendency to hide the growing distress within the bank by reporting lower LLPs, thus effectively inflating reported earnings. In the next stage, we extract the predicted estimates of $DLLPs$, conditional on the bank facing increasing default risk. We then examine the association between predicted *DLLPs* and future stock price crash risk. The results are presented in Table 9 (Panel B). Consistent with prior studies on the valuation effects of bank discretionary *LLPs*, we find that stock price is negatively associated with the level of predicted *DLLPs*[11]. We conjecture that banks which had built up excess LLPs (positive predicted *DLLPs*) in the face of mounting distress were better positioned to withstand a hit from lower future earnings and, thus, were associated with lower stock price crash risk. Moreover, higher reported *LLPs* may have elicited a positive response from the stock market for such banks. On the other hand, banks that had lower-than-expected *LLPs* (negative predicted *DLLPs*) may have not been adequately prepared to withstand future distress and were thus associated with higher crash risk. Our results may indicate a managerial tendency to withhold information on growing default risk by managing discretionary *LLPs*, which in turn results in higher stock price crash risk.

7. Conclusion

Banks are the lifeblood of an economy. An isolated bank default may spill over to connected institutions across the global financial system, potentially inducing distress even in otherwise healthy banks. The GFC is a stark example of how default by a few institutions may have a domino effect across the entire system. Given the systemic importance of these financial institutions, it is imperative for regulators to closely monitor the performance of banks and ensure timely recognition of signs of financial distress. This becomes challenging because banks' financial statements are inherently opaque and difficult to value, unlike non-financial firms. A major portion of bank balance sheets are composed of loans, most of which are privately negotiated with borrowers. Bank management possesses private information on these loans' true value, which exacerbates information asymmetry with outside stakeholders such as

[11] In Table 9, we have used Opacity, and not the three-month change in Opacity as the variable. The results remain similar when we use three-month change in Opacity as in Table 8.

investors. Loan loss provisions *(LLPs*) are a channel through which investors may build an idea of the present financial health of the bank and future earnings prospects. This means timely recognition of LLPs helps create transparency about the bank's financial health. *LLPs,* however, are highly discretionary in nature. Bank managers exercise considerable judgment over the actual amount of *LLPs* that are reported to the public. Motivated by personal wealth and/or career concerns, managers may choose to withhold unfavourable news about a bank's financial health and thus report higher or lower-than-expected *LLPs*. This exacerbates information asymmetry between bank management and external shareholders and potentially draws adverse reactions from the stock market – a price crash - when the bad news is ultimately revealed. This study examines whether a sudden increase in bank distress affects future stock price crash risk. We examine a bank-month panel of US publicly listed commercial banks over the period 2000-2022. We find that a three-month increase in the probability of default is significantly associated with future stock price crash risk. Our results withstand multiple robustness tests, including alternative measures of crash risk and default risk, tests of reverse causality, and functional form misspecification tests. We further document empirical evidence that an increase in default probability is associated with lower discretionary *LLPs* in the next period, which may indicate managerial attempts at obscuring the mounting financial weakness of the bank from external shareholders. Distressed banks that show lower *LLPs* in the next period are also associated with a higher likelihood of future stock price crash. We also find that financially distressed banks that build up provisions by higher discretionary *LLPs* benefit from a lower likelihood of a stock price crash. This may be on account of a safety buffer from excess provisions, which draws a favourable stock market response. Our study emphasizes the importance of supervising and monitoring distressed banks from an early stage. Such banks are especially susceptible to the management of *LLPs* and bad news hoarding, which ultimately culminates in a stock price crash risk once the distress is revealed.

References

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

Andreou, C.K., Andreou, P.C., and Lambertides, N. (2021). Financial Distress Risk and Stock Price Crashes. *Journal of Corporate Finance 67, 101870.*

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

Barth, M. E., Gómez-Biscarri, J., Kasznik, R., & López-Espinosa, G. (2012). Fair value accounting, earnings management and the use of available-for-sale instruments by bank managers.

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

Beaver, W.H. and Engel, E.E. (1996). Discretionary behaviour with respect to allowances for loan losses and the behaviour of security prices. *Journal of Accounting and Economics 22(1-3), 177-206*

Bharath, S. and Shumway, T. (2008). Forecasting Default with the Merton Distance to Default Model. *Review of Financial Studies 21-3, 1339-1369.*

Blau, B.M., Brough, T.J. and Griffith, T.G. (2017). Bank opacity and the efficiency of stock prices. *Journal of Banking and Finance 76, 32-47*

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

Bushman, R. M., and C. D. Williams. 2012. Accounting discretion, loan loss provisioning, and discipline of banks' risk-taking. *Journal of Accounting and Economics 54:1–18.*

Callen, J. L., & Fang, X. (2017). Crash risk and the auditor–client relationship. *Contemporary Accounting Research*, *34*(3), 1715-1750.

Campbell, T.S. and Kracaw, W.A. (1980). Information production, market signalling, and the theory of financial intermediation. *The Journal of Finance 35(4), 863-882*

Chang, X., Chen, Y., and Zolotoy, L. (2017). Stock Liquidity and stock price crash risk. *Journal of Financial and Quantitative Analysis 52(4), 1605-1637*

Chan-Lau, J.A. and Sy, A.N.R. (2007). Distance-to-default in banking: A bridge too far? *Journal of Banking Regulation 9, 14-24.*

Chava, S., & Jarrow, R. A. (2004). Bankruptcy prediction with industry effects. *Review of finance*, *8*(4), 537-569.

Chen, J., Hong, H., and 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, L. J., Cornett, M. M., Marcus, A. J., & Tehranian, H. (2014). Bank earnings management and tail risk during the financial crisis. *Journal of Money, Credit and Banking*, *46*(1), 171-197.

Dechow, P., W, Ge., and C. Shrand (2010). Understanding Earnings Quality: A Review of the Proxies, their determinants, and their consequences. *Journal of Accounting and Economics 50: 344-401*

Dechow, P.M., Sloan, R.G. and Sweeney, A.P. (1995). Detecting Earnings Management. *The Accounting Review 70(2), 193-225.*

Dewally, M. and Shao, Y. (2013). Financial Derivatives, opacity, and crash risk: Evidence from large US banks. *Journal of Financial Stability 9, 565-577.*

Doan, A. T., Lin, K. L., & Shen, C. H. (2024). Bank ownership concentration and stock price informativeness: International evidence. *European Financial Management*, *30*(1), 92-128.

Financial Accounting Standards Board. (2012). Proposed accounting standards update, financial instruments e credit losses. Financial Accounting Series,
Exposure Draft, Subtopic, 825e915 (December 20, 2012).

Harada, K., Ito, T., and Takahashi, S. (2013). Is the distance to default a good measure in predicting bank failures? A case study of Japanese major banks. *Japan and the World Economy 27, 70-82*

Healy, P.M. (1985). The Effect of Bonus Schemes on Accounting Decisions. *Journal of Accounting and Economics 7, 85-107.*

Hedge, S. and Kozlowski, S. (2021). Discretionary loan loss provisioning and bank stock returns: The role of economic booms and busts. *Journal of Banking and Finance 130, 106186*

Hoelscher, D. and Quintyn, M. (2003). Managing systemic banking crises. *Occasional Paper No. 224, International Monetary Fund, Washington.*

Hong, H., Stein, J.C., 2003. Differences of opinion, short-sales constraints, and market crashes. *Review of Financial Studies 16 (2), 487–525.*

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

Islam, M.S., Alam, M.S., Hasan, S.B., and Mollah, S. (2022). Firm-level political risk and distance-to-default. *Journal of Financial Stability 63, 101082*

Jensen, M. and Meckling, W. (1976). Theory of the firm: Managerial behaviour, agency costs, and ownership structure. *Journal of Financial Economics 3(4), 305-360*

Jiang, L., Levine, R., and Lin, C. (2016). Competition and Bank Opacity. *The Review of Financial Studies 29(7), 1911-1942.*

Jin, L. and Myers, S.C. (2006) R2 around the world: New theory and new tests. *Journal of Financial Economics, Volume 79, Issue 2, Pages 257-292, ISSN 0304-405X.*

Jones, J. (1991). Earnings Management during Import Relief Investigations. *Journal of Accounting Research 29, 193-228.*

Jung, T., Kim, N., Kim, Y., and Na, H.J. (2019). Bad news withholding and stock price crash risk of banks. *Asia-Pacific Journal of Financial Studies 48(6), 777-807*

Kanagaretnam, K., Lobo, G., and Yang, D. (2005). Determinants of signalling by banks through loan loss provisions. *Journal of Business Research 58(3), 312-320*

Kanagaretnam, K., Lobo, G.J., and Mathieu, R. (2003). Managerial incentives for income smoothing through bank loan loss provisions. *Review of Quantitative Finance and Accounting 20, 63-80*

Kang, M.J., Kim, Y.H., and Liao, Q. (2019). Do bankers on the board reduce crash risk? *European Financial Management 26, 684-723.*

Kilic, E., Lobo, G. J., Ranasinghe, T., & Sivaramakrishnan, K. (2013). The impact of SFAS 133 on income smoothing by banks through loan loss provisions. *The Accounting Review*, *88*(1), 233-260.

Kim, J. and Zhang, L. (2016). Accounting conservatism and stock price crash risk: Firm-level evidence. *Contemporary Accounting Research 33(1), 412-441*

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

Kim, J.B., Wang, Z., and Zhang, L. (2016). CEO Overconfidence and stock price crash risk. *Contemporary Accounting Research 33(4), 1720-1749.*

Laeven, L. and Majnoni, G. (2003). Loan loss provisioning and economic slowdowns: too much, too late? *Journal of Financial Intermediation 12(2), 178-197*

Liu, C. C., Ryan, S. G., & Wahlen, J. M. (1997). Differential valuation implications of loan loss provisions across banks and fiscal quarters. *Accounting Review*, 133-146.

Lobo, G.J. and Yang, D.H. (2001). Bank managers' heterogenous decisions on discretionary loan loss provisions. *Review of Quantitative Finance and Accounting 16, 223-250*

Lucas, D. (2019). Measuring the cost of bailouts. *Annual Review of Financial Economics 11: 85-108*

Ma, CK (1988). Loan loss reserved and income smoothing: the experience in the US banking industry. *Journal of Business Finance & Accounting 15(4), 487-497.*

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

Morgan, D.P. (2002). Rating banks: Risk and uncertainty in an opaque industry. *American Economic Review 92(4), 874-888*

Morris, R., Kang, H., and Jie, J. (2016). The determinants and value relevance of banks' discretionary loan loss provisions during the financial crisis. *Journal of Contemporary Accounting and Economics 12(2), 176-190*

Nagel, S. and Purnanandam, A. (2020). Banks' Risk Dynamics and Distance to Default. *Review of Financial Studies 33(6), 2421-2467.*

Schuermann, T., & Stiroh, K. J. (2006). *Visible and hidden risk factors for banks* (No. 252). Staff Report.

Wahlen, J.M. (1994). The nature of information in commercial bank loan loss disclosures. *The Accounting Review 69(3), 455-478*

Zheng, Y. (2020). Does Bank Opacity affect Lending? *Journal of Banking and Finance 119, 105900, ISSN 0378-4266.*

Appendix 1: Variable Descriptions and Sources of Data

Variable	Definition	Data Source(s)
Crash Measures		
c309	indicator variable for crash risk, measured as 1 if within the month the bank experiences at least one firm-specific weekly return 3.09 standard deviations below the mean return for the estimation period	Daily stock price data from Compustat NA Daily Files; Authors' calculations
c320	indicator variable for crash risk, measured as 1 if within the month the bank experiences at least one firm-specific weekly return 3.20 standard deviations below the mean return for the estimation period	Daily stock price data from Compustat NA Daily Files; Authors' calculations
NCSKEW	Negative conditional skewness; estimated for each month as per eq.3 using firm-specific weekly returns as estimated by eq.2	Daily stock price data from Compustat NA Daily Files; Authors' calculations
DUVOL	Down-to-up volatility, estimated for each month as per eq.4 using firm-specific weekly returns computed using eq.2	Daily stock price data from Compustat NA Daily Files; Authors' calculations

Default variables and components		
D	Total debt, calcluated as the sum of (a) short term debt, (b) long term debt, (c) total deposits, and (d) preferred equity	Compustat NA Quarterly Files, S&P Capital IQ
FV	Total firm value, calculated as the sum of total debt (D) and equity market capitalization	Authors' calculations
AR	Cumulative 12-month rolling equity returns	Authors' calculations
σ_{AS}	Asset volatility, which is the standard deviation of firm assets. It is estimated as the weighted average standard deviation of equity and debt. Standard deviation of equity, σ_E, is estimated as the standard deviation of	Authors' calculations

	equity returns over prior 12-month window. Standard deviation of debt, σ_D, is $0.05+0.25\sigma_E$.	
DD	Distance to default estimated as per eq.5, as in Bharath and Shumway (2008)	Authors' calculations
PD	Probability of default estimated as per eq.6	Authors' calculations
ΔPD	Three-month change in probability of default, estimated as per eq.7	Authors' calculations
Control Variables		
Size	Natural logarithm of Total Assets	Compustat NA Quarterly Files; Authors' calculations
ROA	Net Income to Total Assets ratio	Compustat NA Quarterly Files; Authors' calculations
LEV	Total Equity to Total Assets ratio	Compustat NA Quarterly Files; Authors' calculations
MB	Ratio of equity market capitalization to book value of common equity	Compustat NA Quarterly Files; Authors' calculations
RET	3-Month cumulative equity returns	Authors' calculations
DTURN	Detrended equity turnover, calculated as monthly turnover minus average monthly turnover over the prior 12-month period	Authors' calculations
HHI	Herfindahl-Hirschman Index, calculated as the squared of each bank's market share of assets, i.e. bank's assets divided by total industry assets in USD millions	Authors' calculations
NIM	Net interest margin, in percentage points	S&P Capital IQ
NPL	Non-performing loans to total loans ratio, in percentage points	S&P Capital IQ
MDZ	Market data-based z-score, estimated as per Lepetit et al. (2007)	Authors' calculations
Opacity1	Measure of opacity estimated as the natural log of residuals obtained from estimation of eq. 8, following Jiang et al. (2016)	Authors' calculations
Opacity2	Measure of opacity estimated as the natural log of residuals obtained from estimation of eq. 9, following Zheng (2020)	Authors' calculations