Essays on Systemic Risk and Financial Stability

Inauguraldissertation

zur

Erlangung des Doktorgrades

der

Wirtschafts- und Sozialwissenschaftlichen Fakultät

der

Universität zu Köln

2020

vorgelegt

von

Arndt-Gerrit Kund, M.Sc. & MIM

aus

Ahlen

Referent: Univ.-Prof. Dr. Thomas Hartmann-Wendels

Korreferent: Univ.-Prof. Dr. Dieter Hess

Tag der Promotion: 26. Mai 2021

Acknowledgments

I want to thank Prof. Dr. Thomas Hartmann-Wendels for not only giving me the opportunity to write this thesis, but also for allowing me to pursue my research interests and to support them wherever possible. His personal guidance and expertise were valuable assets throughout my doctoral studies, and I am thankful that he enabled me to discuss my ideas at multiple conferences around the globe with colleagues. They critically challenged my thinking, and had interesting points of view, which benefited this thesis substantially. Furthermore, I want to thank Prof. Dr. Dieter Hess for serving as co-referee and Prof. Dr. Heinrich Schradin for chairing the dissertation committee.

I am grateful for my colleagues at the University of Cologne, with whom I had not only a pleasant time on and off campus, but also vivid and interesting discussions. Amongst them, my co-authors Daniel Rugilo and Matthias Petras stand out. They were not only superb collaborators, but also genuine friends, which made the writing of this dissertation a great pleasure. I am humbled by having had the opportunity to work with Prof. Dr. David Lando during my research stay at the Copenhagen Business School. He provided valuable comments and an insightful environment, from which I learned a lot.

This thesis is not only the sum of many professional experiences, but also the people around me. I am especially thankful for my girlfriend, who always had my back, and was there, when I needed her the most. Likewise, knowing that I could always count on the unconditional support of my parents and brother has been a great relief. Please accept these words of gratitude as an honest appreciation for everything you have done for me.

Arndt-Gerrit Kund
October 2020

This dissertation consists of three self-containing essays:

Kund (2018): Can Systemic Risk Measures explain Bank Defaults?

05/2018	35th Annual Conference of the French Finance Association (AFFI)
06/2018	35th International Symposium on Money, Banking and Finance (GdRe)
07/2018	25th Global Finance Conference (GFC)
11/2018	58th Meeting of the Southern Finance Association (SFA)

Kund and Petras (2019): Can CoCo-bonds mitigate Systemic Risk?

11/2019	7th Workshop in International Finance & Economics (LaReFi)*
03/2020	82nd Meeting of the German Association for Business Research (VHB)*
04/2020	56th Meeting of the Eastern Finance Association (EFA)$^{\circ}$
08/2020	69th Meeting of the Midwest Finance Association (MFA)

Kund and Rugilo (2018): Does IFRS 9 increase Financial Stability?

04/2019	7th Research Retreat of the Center for Financial Frictions (FRIC)$^{\diamond}$
05/2019	9th International Conference of Financial Engineering & Banking (FEBS)
06/2019	26th Global Finance Conference (GFC)
06/2019	6th Conference on International Finance (INFINITI)
06/2019	6th Young Finance Scholars Conference (JBF)
06/2019	36th Annual Conference of the French Finance Association (AFFI)
06/2019	8th Conference on Accounting & Regulation (EIASM)*
06/2019	20th Meeting of the European Financial Management Association (EFMA)
10/2019	49th Meeting of the Financial Management Association (FMA)$^{\diamond}$
11/2019	59th Meeting of the Southern Finance Association (SFA)

$*$ Attended by co-author.

$\diamond$ Attendance financially supported by the German Academic Exchange Service (DAAD).

$\circ$ Postponed due to COVID-19.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

"You never want a serious crisis to go to waste."

– Rahm Emanuel (former White House Chief of Staff)

The Subprime Crisis led to many consequences not only for the real economy, but also the banking sector in particular. More than ten years later, we can now observe how the lessons learned have found their way into the current regulation. While some countries have enacted entirely new legislation (e.g. Dodd-Frank Act), others have decided to incorporate recent developments of the Basel Accords into local law (i.e. Capital Requirements Regulation). In line with Rahm Emanuel's quotation, this dissertation looks at these changes and investigates in three self-contained essays, whether we have indeed learned from the 2008 financial crisis.

Out of the multitude of research questions that arise from the fundamentally changed regulatory landscape, my analyses focus on two branches in particular: systemic risk and impairment models. The first and second essay discuss the measurement of systemic risk and investigate its consequences in the context of bank resilience. In the third essay, the impact of the revised calculation of credit losses on financial stability is discussed.

Systemic risk has received widespread attention during the financial crisis of 2008 and been thoroughly studied since. However, despite the multitude of research published on the topic, many questions remain unanswered. To begin with, the absence of a common definition of systemic risk highlights the need for further research in unparalleled ways. The ECB (2010a) illustrates the complexity of a uniform understanding of systemic risk by deriving a "systemic risk cube", in which every dimension grasps a different aspect of systemic risk. If one is to apply this classification to the multitude of systemic risk measures (SRM), it becomes apparent that no measure accounts for all dimensions of systemic risk. Hence, the question arises, which of the many SRM possesses the most explanatory power in terms of systemic risk.

I address this research question in the first essay by using the Uniform Bank Performance Reports (UBPR). Doing so allows me to construct a panel in which I observe quarterly balance sheet and profit and loss (P&L) information of 22,751 banks in the United States from 1980 until 2013. In combination with the "Failed Bank List" from the Federal Deposit Insurance Corporation (FDIC), I obtain a data set in which I can relate the failure of 2,044 banks to their fundamental data. Using this sample has four distinct advantages. First, I observe many defaults, because the sample covers 34 years. Consequently, the results are not solely based on a few outliers, but different banks throughout the observation period. Second, the long time frame allows me to incorporate different financial crises, so that the applied methodology does not suffer from being overly exposed to a specific banking or market crises (see Berger and Bouwman (2013)). Third, as all observations are U.S. banks, they are subject to the same accounting standards and regulations, which eliminates potential immeasurable noise. Fourth, the broad range of bank business models and sizes remedies possible limitations such as in Kolari et al. (2002). Taken together, the sample allows to generate meaningful inference in the context of systemic risk, while yielding a high degree of generalization.

The analysis consists of two consecutive steps. I initially derive a default prediction model, which only consists of idiosyncratic risk measures. Applying this model should only identify bank failures that can be related to the fundamental data of the respective bank. If the considered SRM are true measures of systemic risk, they should increase the forecasting accuracy of the first model, by also being able to account for systemically induced bank defaults. Hence, I posit the hypothesis, that the addition of SRM improves the predictability of bank failures. I test this hypothesis for the most established SRM, namely ΔCoVaR, MES, and SRISK. My results show that ΔCoVaR does not benefit the identification of bankruptcy in financial institutions. While MES and SRISK yield more accurate forecasts of bank defaults, the robustness tests indicate that only SRISK does so consistently.

The results are intriguing, because they entail profound implications. On the one hand, they give policy makers and regulators concrete evidence, which SRM to use and hence aid the regulator triage. On the other hand, they highlight the harsh reality of measuring systemic risk with all its deficits and motivates further research into the topic.

I pursue this challenge together with my co-author Matthias Petras and extend the understanding of systemic risk, in the second essay of this thesis. We shed light on a growing body of the literature, which concerns the usage of hybrid capital in banks. In particular we look to contingent convertible bonds (CoCo-bonds) as a subset of hybrid capital. It has seen stellar growth after the 2008 financial crisis, because it combines the advantages of debt and equity in one financial instrument, while allowing banks to fulfill their Pillar I requirements in line with the Basel Accords. Time and again, CoCo-bonds have been praised not only due to the aforementioned benefits, but also as a tool to increase the systemic resilience of banks. However, empirical evidence for this attribution is sparse (Avdjiev et al. (2013)). We extend the literature by empirically testing the hypothesis whether CoCo-bonds reduce the systemic risk of the issuing banks. Taking the results of the first essay into consideration, we compute SRISK as a measure of systemic risk for 126 banks from 33 countries around the world. Our data set covers many banks that have issued multiple CoCo-bonds over the analyzed time frame from 2012 until 2018. None of them has been called or triggered, such that our sample is free of a possible survivor bias.

Our initial results suggest that CoCo-bonds do not reduce systemic risk. This finding is puzzling, given that CoCo-bonds create de facto additional loss-absorbing capital for the issuing bank. We relate this observation to a transmission mechanism between accounting law and CoCo-bonds that is likely to distort the risk reduction of CoCo-bonds. Hence, we suggest an alternative calculation to address this issue. Subsequent analyses show that CoCo-bonds reduce systemic risk, irrespective of their balance sheet treatment. Through different robustness tests, we reinstate the validity of our solution for different parametrizations of SRISK (i.e. k and $LRMES$), issuance effects and generate evidence against possible endogeneity concerns. Our results entail both theoretical, and practical implications. They substantiate the usage of CoCo-bonds as means to increase the systemic resilience of issuing banks from the regulator's perspective. At the same time, our findings corroborate the usage of CoCo-bonds from the bank management's point of view, which should continue to reap the benefits of combining the advantages of debt and equity.

The third essay demonstrates further unintended consequences of accounting law at the intersection of financial stability. I proceed to analyze them in more detail with my co-author Daniel Rugilo. The starting point of our analysis is the Subprime Crisis of 2008. It highlighted severe deficits in the incurred loss accounting of IAS 39, which was consequently replaced with the expected loss accounting of IFRS 9. It intends to anticipate credit losses, in an attempt to reduce volatility in the financial sector, and hence foster its stability. However, the implementation of IFRS 9 has released two opposing forces in the context of financial stability. While the expected loss accounting reduces jumps in impairments that were induced by the timely disparity under the previous accounting standard (i.e. "cliff-effect"), it potentially depletes banks' capital levels through the excessive front-loading of impairments. We thus set out to investigate the net impact of the new accounting standard in the context of financial stability by postulating three hypotheses. While hypothesis one and two highlight the (dis-)advantages, hypothesis three constitutes the synopsis, which concludes on the net effect. In particular, we investigate, whether IFRS 9 (i) reduces the volatility of impairments, (ii) erodes the capital base of banks, and (iii) reduces bank stability in particular.

Given that IFRS 9 has only been enacted from January, 1st 2018 forth, data availability poses a central challenge to our research question. We remedy the absence of archival data by reverting to the European bank stress tests, which include forecasts until 2020. By analyzing the data from 2014 forth, we obtain a panel of 43 banks from 15 different European countries, which report in accordance with IAS 39 as well as IFRS 9 such that we can derive meaningful inference from contrasting the two time series. Using the European stress test results does not only remedy the data constraint, but also benefits our identification strategy as it entails additional favorable properties. First, it is conducted as a constrained bottom-up exercise, such that all banks conduct their own simulations in accordance with the principles laid out by the European Commission (EC), as well as the European Systemic Risk Board (ESRB). As a result, we can eliminate noise from the macroeconomic environment by exhaustively controlling for it. Second, the stress test is conducted under the static balance sheet assumption, which replaces maturing assets and liabilities with similar financial instruments. We can thus exclude immeasurable externalities, allowing us to assess the true impact of IFRS 9. Third, the stress test is conducted not only for a baseline, but also an adverse scenario, in which the economy further deteriorates. The comparison between the two scenarios allows us to derive meaningful insights on how the respective accounting standards contribute to procyclicality.

In accordance with our first and second hypothesis, we find that IFRS 9 reduces the volatility of impairments, while doing so at the expense of impeding the potential to increase the capital base by retaining earnings. As the results are of diametrically opposing influence on financial stability, we proceed with our third hypothesis, which investigates the net influence on bank stability. We use the z-Score of Goetz (2018) as a proxy for bank resilience and show that IFRS 9 exerts a stronger influence in the baseline scenario, which consequently narrows the gap to the adverse scenario. It thus appears as if loans become more expensive due to earlier impairments. At the same time, the resilience of banks increases, as they do not suffer from steep increases in their likelihood of failure during economic downturns. We thus conclude that the identified concerns of IAS 39 have been addressed.

Coming back to Rahm Emanuel's encouragement to learn from financial crises, this thesis yields an ambiguous answer. The first two essays demonstrate deficits in the currently employed systemic risk measures, which (i) are not strictly generic measures of systemic distress, and (ii) do not properly account for hybrid capital, respectively CoCo-bonds in particular. At the same time the third essay shows that newly implemented IFRS 9 expected credit loss model systematically addresses identified shortcomings of the previous accounting standard. With the upcoming implementation of a similar model in the U.S., it appears as if this source of systemic fragility has been resolved.

Taken together, this thesis contributes to the research area by dissecting the shortcomings of systemic risk measures, and subsequently raising new research questions. Although a multitude of techniques has been proposed since the 2008 financial crisis, they are each subject to individual (dis-)advantages. No approach has emerged superior over all dimensions of systemic risk, such that further research into the topic is warranted. My results thus urge policy makers and regulators alike to not become complacent in measuring systemic risk. To the contrary, given the novelty of systemic risk measures, additional research is needed in order to mature this strand of the literature. A promising endeavor could be the incorporation of hybrid capital, which grows in prevalence, yet remains not properly accounted for. Our findings regarding the stability enhancing effects of IFRS 9 should urge U.S. regulators to implement the equivalent CECL regulation without undue delay.

Chapter 2

Can Systemic Risk Measures explain Bank Defaults?

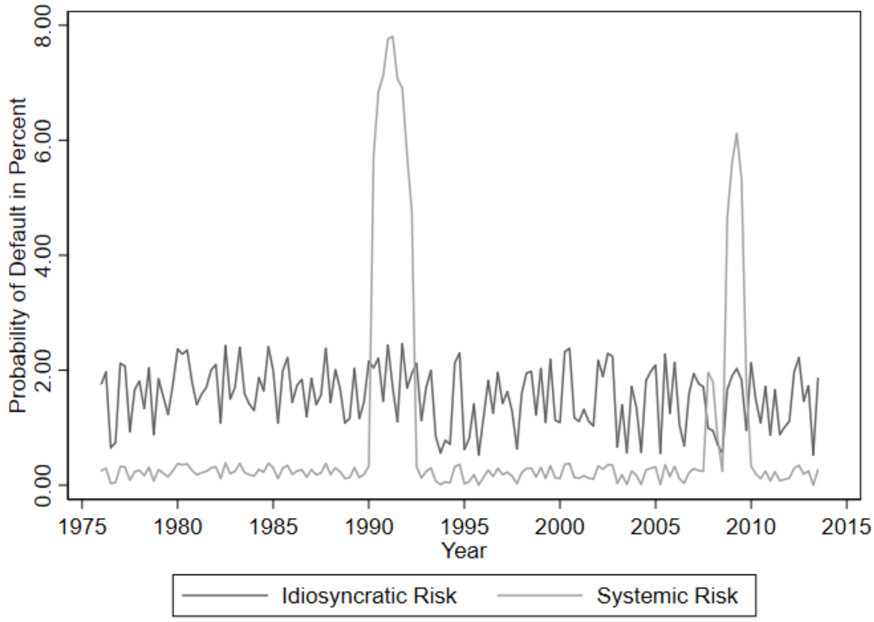
2.1 Introduction

Financial crises and the subsequent bank defaults continue to impose significant costs on the real economy (Hoelscher and Quintyn (2003); Reinhart and Rogoff (2014); Laeven and Valencia (2018)). Reducing these costs is a beneficial undertaking for all stakeholders of the economy. While banks must be allowed to fail, in order to prevent moral hazard as illustrated by Korte (2014), there are times, when saving them is desirable. Examples are given by Friedman and Schwartz (1963) who show that many banks during the Great Depression were only illiquid, but not insolvent, and should thus have been saved. Hence, the identification of failing banks ex ante is of paramount importance in order to ensure an efficient discrimination between viable and non-viable banks. As Homar and van Wijnbergen (2017) show, doing so allows the latter to be recapitalized, which in turn aids the economic recovery.

This paper embraces the challenge of default prediction and approaches bank failures from a new perspective. While the literature on the topic is plentiful, it shares one central assumption: in line with the groundbreaking work of Altman (1968) it is presumed that bank defaults can be explained by fundamental data as shown in Calomiris and Mason (2003). Thus, most of the developed models draw excessively from idiosyncratic measures, which they also use to explain different financial crises (see Berger and Bouwman (2013)). However, as King et al. (2006) and Weiß et al. (2014) show, the drivers of systemic risk are unique for each crisis. Consequently, the number of needed variables grows at least linearly with the frequency of crises. This dependence exposes such models to overfitting, which I remedy by using generic measures of systemic distress. Systemic risk measures (SRM) account for different financial crises irrespective of their source and hence allow me to decouple the number of variables from the frequency of financial crises. As a result, I derive a parsimonious but holistic model of bank default prediction, which combines different risk sources, while keeping the number of input factors constant.

The underlying rational is intuitive: it assumes that bank defaults can be related to either idiosyncratic or systemic risk. Both can grow independently of another, but jointly constitute the aggregated default risk, as illustrated in Figure (2.1). In line with the seminal work of Black and Cox (1976) bankruptcy is triggered when at least one of the risk factors exceeds the default boundary.

Figure 2.1: Constituents of the default risk.



Idiosyncratic risk can be directly related to an institution, and hence be explained by a combination of its fundamental data (i.e. its balance sheet, as well as profit and loss information). Although Acharya and Ryan (2016) emphasize the necessity of high accounting quality in order for this assumption to hold, it is often made, as i.a. in microprudential regulation, which is based on fundamental data, too. In contrast, systemic risk is harder to conceptualize. This paper understands it as the risk that all institutions within a system are subject to, owed to the design of the system. Systemic events are self-enforcing endogenous feedback loops in this system. Brunnermeier and Pedersen (2008) describe such a system, were initial losses lead to the sale of securities, which deteriorate in price, and hence necessitate further security sales to cover new losses. Chen et al. (2016) document instances where the cross-holding of assets or liabilities (i.e. the network channel), or the competition for scant funding (i.e. the liquidity channel) amplifies downturns in the system. The risk that such events materialize is the actual systemic risk, as described by Daníelsson et al. (2013).

This paper goes beyond the current literature by making two main contributions. First, the literature on the prediction of bank defaults is connected to measures of systemic risk.

In doing so, it becomes possible to derive a holistic model of bank failure, which accounts for insights from both streams of the literature. Second, this new model is evaluated and benchmarked against different measures of systemic risk. In doing so, insights on the explanatory power of individual SRM for identifying distressed banks are derived. As a result, regulatory triage is facilitated by giving supervisors concrete evidence, which SRM to use. In particular, I show that SRISK appears to be the most informative of the tested SRM.

The structure of this paper is as follows: Section (2.2) revises means of default prediction for banks, and introduces the SRM, which are applied in this paper. Next, the data set is described, and additional metrics for the pursuant calculations are derived in Section (2.3). The applied model is outlined in more detail in the following fourth section. Section (2.5) presents the results, which are subject to a plurality of robustness tests in the ensuing sixth section. The final section concludes and gives an outlook into possible research questions arising from this work.

2.2 Theoretical Background

2.2.1 Predicting Bank Defaults

Predicting bank defaults is a tedious task for a plurality of reasons. Unlike other companies, banks can net their exposures under certain accounting regimes, such as U.S. GAAP (ASC 210-20), on which this data set is based. This property leads to a distorted view on the actual exposures a bank may have. It is especially problematic when defaults occur, and the theory of offsetting payments has to be rejected. The implications of this approach become evident, when looking at the balance sheet of Deutsche Bank, which reports under both, U.S. GAAP and IFRS. While its total assets were EUR 2,282 billion under IFRS, the application of netting in line with U.S. GAAP has brought it down to EUR 1,296 billion in the third quarter of 2011 (Deutsche Bank (2012)). This difference can have significant implications for capital adequacy ratios, which can increase by up to 50.00 % as Hoenig

(2014) comments. Consequently, they shed little light on the de facto health of an institution, and can even be misleading under financial distress as Jungherr (2018) points out.

Moreover, the extend of maturity transformation can induce a significant imparity between assets and liabilities. It constitutes a concealed default risk that materializes as illiquidity, when the required short term funding becomes constrained. In line with this thought, Friedman and Schwartz (1963) argue that many bank failures during the Great Depression could have been averted, had the government addressed these liquidity constraints.

Furthermore, banks have steep default paths, leading to sudden demises that are difficult to foresee, as Vazza and Kraemer (2018) illustrate. Although Martin (1977) points out that convalescence can occur, it is observed seldom. Cole and Gunther (1995) add that the demise occurs just as quickly in small and large banks. They also emphasize that variables, which indicate bank resilience, cannot be interpreted bidirectionally. If they yield poor readings, it must not be an indicator of imminent bank failure. This caveat further complicates the quest for variables that explain bank failures. In response to these challenges, a multitude of techniques have been suggested to assess the resilience of banks. This topic is discussed comprehensively in the work of King et al. (2006), Demyanyk and Hasan (2010) and Giesecke and Kim (2011).

Despite the numerous obstacles in predicting bank failure, a vast strand of literature on the topic exists. It can broadly be divided in parametric (e.g. Merton-type models, logistic regression) and non-parametric (e.g. trait recognition) models. One of the first authors to contribute to the literature on parametric approaches was Altman (1968). He computed a set of financial ratios, based on which a discriminant analysis was conducted. Martin (1977) generalized this approach by employing a logit regression as means of an early warning indicator for commercial bank failure. Since then, a multitude of noteworthy contributions have been made. A recent example is the work of Tong (2015), who uses a logistic regression on a smaller and more aggregated data set. He derives the PD by inferring from the logistic regression function after standardizing the regressors to be normally distributed. As in Zaghdoudi (2013), the used regressors are derived by taking the discriminant analysis of Altman (1968) into account. In accordance with it, Tong finds that certain idiosyncratic ratios explain defaults outstandingly well. In the context of banks, this observation is especially true for the return on equity. By plotting the Bayesian Information Criterion (BIC), relative to the number of regressors, Tong assesses the predictive power of his variables. While the function is initially diminishing, it converges towards a constant and does not appear to be convex. This observation suggests that no minimum has been found, which might hint at an omitted variable. As the model only consists of balance sheet information, this finding supports the hypothesis of this paper that the incorporation of SRM further enhances the forecasting accuracy by remedying the omitted variable problem.

Cox and Wang (2014) apply a related methodology. They enhance Altman's discriminant analysis by applying the "leave-one-out estimation" for cross validation. Comparably to the out of sample validations applied in time series data, one observation of the population is left out when estimating the parameters. This procedure is repeated until each observation has been left out once. As a result, there are as many equations as observations. The aggregate error over all these equations can be collected and minimized with regards to the size of the test. By doing so, they increase the likelihood of not overseeing an ailing bank. Despite this favorable property, their model should not be taken as infallible. It is centered around the recent financial crisis and the model specification correlates significantly with the drivers of the crisis. High loan growth and foreclosure rates, especially in real estate, might perform well as predictors for this crisis, but underperform in other crises. Consequently, their study affirms the objective of this paper in combining idiosyncratic and systemic risk measures as a remedy for potential overfitting.

Another strand of the literature concerns non-parametric models. In this context, the work of Kolari et al. (2002) is noteworthy. They apply a trait recognition model and compare its performance to a logit model for large banks in excess of USD 250 million total assets. While both behave well in-sample, the performance of the logit model deteriorates out-of-sample. As a result, they conclude that trait recognition is the superior method in this context. However, trait recognition yields contentious results owed to the underlying partitioning. Based on metrics such as standard deviations, cut off points are derived as to classify the data. The partitioned data is then split in groups on the basis of the previously specified traits. This procedure appears arguably arbitrary, and comes short of possible economic explanations. Beutel et al. (2019) follow a more structured approach, in which they conclude that non-parametric models can only compete in-sample with standard logit models.

Taken together, the ex ante identification of failing banks remains difficult, in spite of valuable contributions from the cited literature. The meaningfulness of the provided variables is often low owed to netting rules under prominent accounting regimes. As a result, the variables become opaque and less informative. Because of steep default paths, bankruptcies materialize quickly, which makes them difficult to foresee. Addressing these issues has yet to yield a satisfactory model. Despite performing well, the presented models are subject to individual shortcomings and limitations. Tong (2015) owes a clarification with regards to the convexity of his maximum likelihood function and alternates the regressors in his model over time. Cox and Wang (2014) produce a sound model, which however is overly exposed to the drivers of the 2008 financial crisis and thus questionably specified. Kolari et al. (2002) bring forward a robust model for banks in excess of USD 250 million, which is yet to be validated for smaller banks.

2.2.2 Measuring Systemic Risk

While the previous section gave an overview on default prediction models, this section focuses on measuring systemic risk. As per the introduction, this paper understands systemic risk as the risk of distress, which all financial institutions in a given system face due to its design. Because this risk is latent and cannot be measured, unless a systemic event occurs, quantifying systemic risk is intricate and may yield ambiguous results.

Hence, there is no general consensus on measuring systemic risk. Instead, a multitude of different approaches have been postulated, as most recently collected by Benoit et al. (2017). Two major strands of literature can be identified, which differentiate between macro- and microlevel measures of systemic risk. While the first assess the riskiness of the entire financial system, the latter focus on individual banks, and are hence of special interest for the purpose of this paper. As a result, ΔCoVaR, MES, and SRISK, are analyzed in more detail. They were chosen due to their prevalence in the literature, and feasibility of computation in the context of the research question.

In the interest of completeness, it shall be said that the literature proposes even more measures of systemic risk. Noteworthy examples are the work of Billio et al. (2012) and Patro et al. (2013) who assess correlation matrices. Furthermore, measures relying on the spreads of credit default swaps (CDS) are postulated by Huang et al. (2009) and Chan-Lau (2010). However, they are excluded from the analysis because they cannot be computed for the majority of the analyzed institutions owed to the lack of tradable CDS in the institution. A more detailed overview regarding the universe of suggested SRM can be obtained by the work of de Bandt and Hartmann (2000), FSB (2009), in the Financial Stability Reviews of the ECB (2010a,b), respectively Bisias et al. (2012) and Benoit et al. (2017).

Adrian and Brunnermeier (2016) propose ΔCoVaR to measure the marginal contribution of an individual bank to the aggregate systemic risk of the financial sector. In a first step, they define a lagged vector M_{t-1}, which contains proxies for the health of the market. It is then related to the growth rate of total assets of bank at market values of bank i at time t (X_t^i) with a quantile regression. As a result, a vector of estimates is obtained, henceforth denoted as γ^i. α refers to the intercept, whereas the error term is denoted by ϵ.

$$X_t^i = \alpha^i + \gamma^i M_{t-1} + \epsilon_t^i \tag{2.1}$$

By incorporating the bank specific growth rate X_t^i from Equation (2.1) on the right hand side of Equation (2.2), the growth rate of the financial system can be estimated conditional on which bank is denoted by the subscript i.

$$X_t^{system} = \alpha^{system|i} + \beta^{system|i} X_t^i + \gamma^{system|i} M_{t-1} + \epsilon_t^{system|i} \tag{2.2}$$

Using the estimates of γ^i from Equation (2.2), the Value at Risk (VaR) of an institution is predicted to a certain quantile denoted by q, as shown in Equation (2.3).

$$VaR_t^i(q) = \hat{\alpha}_q^i + \hat{\gamma}_q^i M_{t-1} \tag{2.3}$$

In order to predict the values on a market level, the VaR estimate from Equation (2.3) is incorporated on the right hand side of Equation (2.4) as to derive the CoVaR.

$$CoVaR_t^i(q) = \hat{\alpha}^{system|i} + \hat{\beta}^{system|i} VaR_t^i(q) + \hat{\gamma}_q^{system|i} M_{t-1} \tag{2.4}$$

Using the conditional estimate of $\hat{\beta}$ from Equation (2.4), the predicted VaR of two different quantiles are then weighted to derive ΔCoVaR. It is the risk-adjusted downturn return of the analyzed institution, relative to the market. As such, it can be understood as the marginal contribution of the analyzed bank to the aggregate systemic risk.

$$\Delta CoVaR_t^i = \hat{\beta}^{system|i}(VaR_t^i(q) - VaR_t^i(50\%)) \tag{2.5}$$

Acharya et al. (2017) propose another metric that assesses systemic risk: the marginal expected shortfall (MES). It transforms the idiosyncratic expected shortfall (ES) into a systemic metric by computing the average returns of the bank not conditional on the VaR of the bank, but the market. In doing so, the bank's returns are ordered relative to a descending vector of market returns. The required information set to do so is denoted by the vector I. In a next step, the MES can be obtained as the average of the bank returns (r) conditional on the market VaR computed to an ex ante defined quantile (q).

$$MES_q^i = -\mathbb{E}\left[r_{bank} \mid I_{market,q}\right] \tag{2.6}$$

This approach allows the market and bank returns to differ from another. While they are independently distributed, their probability functions are not necessarily identical. As a result, the ES of the bank is obtained relative to the worst outcomes of the market and not the worst outcomes of the bank itself. Consequently, the marginal contribution of the bank in the event of a market downturn can be assessed. This difference is a significant advance over former Merton-Style models, where the capital shortfall of an individual bank was calculated, without taking the market into consideration. To ease comparability with other measures, the sign of MES is flipped, such that a growing value translates to rising systemic distress.

Finally, SRISK, will be discussed. Brownlees and Engle (2016) define it as the expected lack of funding that a bank experiences during an extended market down turn, referred to as the Long Run Marginal Expected Shortfall (LRMES). It is calculated as the expected capital shortfall conditional on the occurrence of a systemic event, denoted by c. They define the materialization of systemic risk as a decline in market prices of at least 10 % over the course of a month. In this context, it can be understood, as yet another extension of the ES with similarities to the previously discussed MES. LRMES is the expected value of returns worse than c, adjusted for individual risk through β, as well as time through $\sqrt{h}$.

$$LRMES_{i,t} = -\sqrt{h}\beta_i \mathbb{E}\left(r_{i,t+1} | r_{m,t+1} < c\right) \tag{2.7}$$

After obtaining the LRMES, it is incorporated in the calculation of SRISK by multiplying one minus LRMES with the equity of bank i at time t ($E_{i,t}$). The term is then weighted in order to account for the regulatory capital fraction k. In accordance with the Basel Accords, Brownlees and Engle (2016) set this ratio at 8 %. Although this assumption is questionable in light of recent regulatory developments, I replicate it in order to facilitate comparability. Likewise, it is worth noting that Brownlees and Engle (2016) assume that debt and equity add up to liabilities. In doing so they violate the parity of the balance sheet, as they omit bank deposits, which can widen the funding gap significantly in case of a bank run. In the interest of replicability, I reapply these assumptions and deduct the previous term from the market valued debt ($D_{i,t}$) times the regulatory capital fraction, which can formally be arranged as:

$$SRISK_{i,t} = kD_{i,t} - (1-k)\, E_{i,t} \left(1 - LRMES_{i,t}\right) \tag{2.8}$$

By doing so, one obtains a SRM that contains balance sheet information (i.e. debt and equity), as well as market information as judged by the returns observed in LRMES. Negative values of SRISK can be understood as the funding gap between debt and equity in the case of a systemic event. Hence, it quantifies the additional capital a bank needs to raise, in order to survive a market downturn with the severity of c.

In conclusion, the most relevant SRM in the context of the research question have been presented. All of them have individual strengths and weaknesses, and should thus be taken with a grain of salt. One of the most common critiques is that they fail to address all dimensions of systemic risk (ECB (2010a)). They may be good at estimating losses on a bank level in the case of a systemic event, yet fail to capture the pursuant contagion that ricochets through the financial system. Another concern, as voiced by the ECB (2010b), relates to the problem of identifying the masked growth of systemic risk. Because imbalances in the financial system only evolve gradually, they go unnoticed until their abrupt manifestation. In this context Adrian and Brunnermeier (2016) mention the volatility paradox: systemic risk tends to build up during times of low volatility, which suggests sound markets, whereas the observation should actually be understood as a call for caution as a systemic event might be about to unfold. Daníelsson (2019) affirms this concern, and argues that systemic risk develops, despite all indicators making us believe otherwise.

2.3 Data Set and Methodology

The analyzed data are a compilation of the Uniform Bank Performance Reports (UBPR) from the Federal Financial Institutions Examination Council (FFIEC), and consist of commercial and saving banks in the U.S. from 1980 until 2013. By merging profit and loss (P&L) information with balance sheet data, a sample of 22,751 banks with observations on a quarterly basis is obtained. In doing so, the CERT key was used as unique identifier as it had the same cardinality in both data sets. Pursuant, the initial population is amended with information from the “Failed Bank List” as published by the Federal Deposit Insurance Corporation (FDIC). Doing so allows concluding on the point of default and to mark the corresponding observation in the sample.

The data set has a plurality of favorable properties. While it is an unbalanced sample, it contains 2,044 bank failures, which ensures the model to be reliably calibrated, and not only depend on a few outliers that defaulted. Moreover, the long time frame of 34 years allows the testing of the SRM for various financial crises. As such, it covers not only banking crises, such as the recent Subprime Crisis, or the Savings and Loan Crisis (S&L) of the 1980s, but also market crises, such as the Russian Debt Crisis (1995/1996) and the bailout of Long-Term Capital Management (1998) (Berger and Bouwman (2013)). Furthermore, it remedies the critique of Acharya et al. (2017), who find that most studies on the topic use insufficiently long time series, which do not correct for the infrequency of systemic events. Another benefit of this data set arises from the fact that all banks are from the same country. As a result, they are subject to a homogeneous set of regulations, such that interference from different accounting or regulatory regimes can be kept to a minimum. Likewise, distortions from the macroeconomic environment are reduced, because shocks to it affect all banks in the sample. Lastly, limitations as in Kolari et al. (2002) are not applicable, as the data set consists of banks of different sizes and business models. Consequently, the results generated in this paper have a high degree of generalization and help elicit the true nature of systemic risk.

Taking the established literature into account, I amend the population with additional metrics. Before they can be computed, I have to disaggregate the reported values to a quarterly basis, as they are reported accumulatively by default. In line with Goetz (2018) I include the percentage of non-performing loans (NPL) as the sum of past due and non accruing loans divided by the total loan volume. Furthermore, the return on assets (ROA) is computed as the ratio of operating income minus operating expenses over total assets. Compared to the return on equity (ROE), this definition accounts for the amount of distributable items that can be paid out to claimants in general, and not only equity holders in particular. It was moreover chosen over the ROE, because total assets cannot become negative, while this assertion is not true for total equity. Negative equity does not necessarily constitute a default criterion under U.S. bankruptcy law, such that cases may occur in which both, equity and earnings may be negative. As a result, ROE may be positive, suggesting a sound bank, whereas the opposite is true. Furthermore, ROE is highly susceptible to changes to the bank's leverage, which I analyze separately. I calculate the balance sheet leverage ratio (LR) as total equity divided by total assets, because this data set precedes the introduction of Tier 1 capital from the Basel Accords. Using the regulatory leverage ratio would thus induce a selection bias, as I exclude financial crises, which occurred before the first Basel Accord. The cost-income-ratio (CI) is derived as operating expenses divided by operating income. Ultimately, the loans to deposits ratio (LTD) is calculated as the quotient of total loans over total deposits. In line with Diamond and Rajan (2000) and Diamond and Rajan (2001) it captures the refinancing pressure that a bank would be subject to in case of a bank run. With the intent to further quantify the resilience of a bank, the degree of income diversification is measured using the income diversity measure (ROID) of Laeven and Levine (2007). They use a modified specification of a Herfindahl-Hirschman Index as shown in Equation (2.9). If interest income and non-interest income are roughly the same, the enumerator of the subtrahend converges towards zero, such that the whole subtrahend becomes zero. The minuend is thus not lessened, such that values closer to one indicate higher degrees of diversification.

$$\text{ROID}_{i,t} = 1 - \left| \frac{\text{Interest Income}_{i,t} - \text{Non-Interest Income}_{i,t}}{\text{Total Operating Income}_{i,t}} \right| \tag{2.9}$$

By measuring the income diversification, the deliberations of Vallascas and Keasey (2012) and Weiß et al. (2014) are taken into account. While Weiß et al. (2014) find evidence that non-interest income does not contribute to systemic risk, Vallascas and Keasey (2012) show that it possesses explanatory power in the context of bank defaults. Hence, it appears only prudent to incorporate non-interest income in light of the conflicting results. Furthermore, Köhler (2015) argues that if earnings are well diversified, interest rate shocks as observed during the S&L Crisis are less devastating to the bank's bottom line. Divergent opinions on this hypothesis exist, as documented in the work of DeYoung and Roland (2001). They find that non-interest income tends to be more volatile, making it prone to evaporate quickly during financial distress. Hence, well-diversified banks may actually be more exposed to refinancing risk. De Jonghe (2010) confirms this finding for European banks, and argues that interest income is the most resilient revenue stream. King et al. (2006) show the growing importance of income diversification by contrasting trends at failed banks before and after 1995. In line with that, DeYoung and Torna (2013) analyze the influence of income from nontraditional (i.e. non-interest income) banking activities on bank resilience and find mixed results regarding their implications for bank stability.

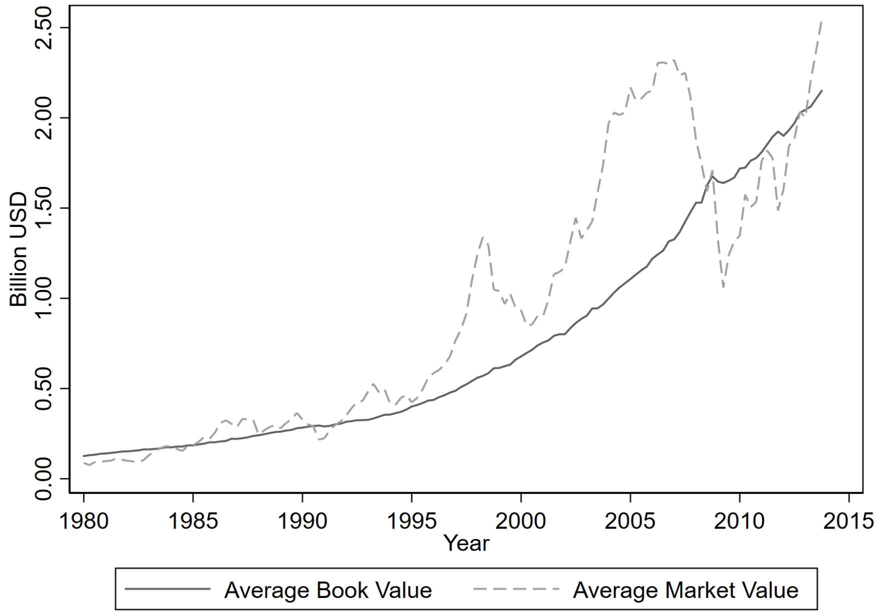
For the implementation of the SRM, it is necessary to obtain market returns. In instances where they were not available, they were derived from the change in approximated market equity values. They were computed by multiplying the book values with a bipartite multiple. It consists of the market to book-ratio for the bank sector, as provided by the Fama-French 48 industries index as a foundation. However, this foundation is only time-variant, but static between banks. Hence, I introduce bank specific variance by adding the compounded annual asset growth rate (CAGR). The intuition for this approach is straight-forward, as it is based on the assumption that investors are willing to pay a premium for banks that outgrow their peers as measured by the CAGR. The CAGR is defined as the quotient of the last reported assets over the first reported assets to the power of one over the observed periods, minus one.

$$\text{CAGR}_i = \left(\frac{\text{assets}_{i,T}}{\text{assets}_{i,t}}\right)^{\frac{1}{T}} - 1 \tag{2.10}$$

$$\text{multiple}_{i,t} = \text{CAGR}_i + \text{Market to Book Ratio}_t \tag{2.11}$$

The estimated multiples are fairly low, as illustrated by the fact that only about 50 % of the data trade at a multiple in excess of one. Owed to the absence of significant intangible values, this characteristic seems reasonable and is in accordance with what can be observed on financial markets. Guerry and Wallmeier (2017) report similar results using an altered version of Tobin's Q, which divides the sum of market-valued equity and book-valued debt by the replacement value, which is the book value of the bank. The accompanying descriptive statistics indicate that their estimates are bound by 0.953 (1.157) on the lower (upper) end, which compares favorably to this data set, where the majority of values concentrates between 0.925 and 1.415. Another rational that attests to the robustness of this transformation stems from bank valuation. Under the view that banks are de facto a well diversified loan portfolio, they would need to trade in excess of their nominal value in order to achieve a multiple above one.

Figure 2.2: Average book value of equity versus average market value of equity.



Despite the reassuring findings, the possibility of a measurement error that would systemically bias the computed multiples has to be discussed. As can be seen in Figure (2.2), the approximated multiples inflate around the occurrence of asset pricing bubbles. This observation gives credibility to the applied transformation, as high multiples coincide with systemic crises as shown by Döring (2016). However, graphic evidence per se should not be taken as infallible. I thus generate numerical evidence in favor of the robustness by regressing the estimated values on the actual values (where available). I find in untabulated results that the differences are marginal, and that the theoretical and empirical values have high reciprocal explanatory power. Similarly, Adrian and Brunnermeier (2016) find that different approaches to deriving market values have little influence on their reported results. It is thus assumed that the influence of the approximations can be neglected for the pursuant calculations.

With regards to the information vector M for the CoVaR model the methodology of Adrian and Brunnermeier (2016) is repeated with minor alternations: the missing data for the S&P 500 Volatility Index (VIX) are interpolated through regressing the VIX on the S&P 100 Volatility Index (VXO) and then predicting the periods in question through the regressors. However, in order to calibrate the regression, the used time frame is expanded to 2015 to account for more recent information. Owed to the lack of access to the three-month repo rate as in the original paper, the difference between the federal funds rate and the three month bill rate is used to approximate the liquidity spread. As a positive side effect of doing so, back-dating the London Interbank Offered Rate (LIBOR) for calibration purposes becomes obsolete. Apart from that, no alternations were made to any assumptions, in the calculation of the used SRM. In order to calculate the necessary returns, the theoretical market values were derived by multiplying the book value of equity with the multiple from Equation (2.11). The quotient of the current theoretical market value over the previous theoretical market value minus one constitutes the estimated return. Table (2.11) in the Appendix contains the descriptive statistics of the used variables, while Table (2.12) yields their correlation coefficients.

2.4 Econometric Model

As argued by Lucas (1976), forecasting outcomes based on relationships observed in historical data can have significant drawbacks, if the underlying patterns are not static. Given the ever changing nature of bank regulation, this critique is warranted and urges caution. Radjan et al. (2010) add in this context that a perfect model of default prediction is utopian because models are not anthropomorphic and fail to account for the incentives of the agents, which may not necessarily be pecuniary. While Cont et al. (2016) underline the importance of monitoring systemic risk, the Subprime Crisis of 2008 has reinstated the concern that models cannot predict defaults flawlessly, irrespective of their sophistication. In line with these deliberations, this paper does not strive to perfectly predict bank defaults, but instead investigates whether the incorporation of SRM yields novel insights in explaining bank defaults.

In doing so, the law of parsimony is applied in order to derive a sparse, but efficient econometric model to forecast bank defaults. It uses nine idiosyncratic variables, and hence differentiates itself from the 25 variables used in Martin (1977) or the 19 used by Cole and

Gunther (1995). This reduction in intricacy relates to the initial deliberation that different crises should not be fitted with idiosyncratic, but generic SRM, as in this paper. In doing so, I follow the argument of Haldane and Madouros (2012) who advocate lean models, as to prevent complexity, which they define as unmanageable risk. Furthermore, the literature has generated evidence in favor of parsimonious specifications. Although Martin (1977) has a broad selection of variables to choose from, it is the most narrow variable selection that produces the best results in his work. Pankoke (2014) attests to this observation, by showing that the simplest models, in the context of systemic risk, tend to yield the highest explanatory power. Likewise, Estrella et al. (2000) show that simple ratios of capitalization are valuable indicators for failing banks. Following these deliberations, this paper stylizes bank defaults ($\hat{D}$) as being triggered by either the idiosyncratic or systemic risk of the preceding period, as shown in Equation (2.12).

$$\underbrace{f(\text{balance sheet}_{i,t}) + g(\text{P\&L}_{i,t})}_{\text{idiosyncratic risk}} + \underbrace{h(\text{SRM}_{i,t})}_{\text{systemic risk}} \implies \hat{D}_{i,t+1} \tag{2.12}$$

As argued in the introduction of this section, idiosyncratic and systemic risk can grow independently of another, but jointly constitute the aggregate default risk, which triggers bankruptcy when it exceeds the default boundary. In this model, the idiosyncratic risk can be related to balance sheet, as well as P&L information. The selection of variables to model the idiosyncratic risk has received broad attention since the early 1980s. Martin (1977) found that capital ratios, as well as liquidity and profitability measures, are good predictors of distress on the bank level. Cole and Gunther (1995) have shown that non-performing loans (NPL), cash, equity, and the cost-income-ratio help explain the riskiness of individual banks. In line with the CAMELS ratios employed by i.a. the Federal Reserve, Cole and Gunther (1998) demonstrate that capital adequacy, asset quality, earnings and liquidity contain significant predictive power. Most recently, Vallascas and Keasey (2012) found non-interest income to be important. While the list of possible variables is long, I do not want to deviate from the preceding discussion of a parsimonious model. At the same time, the choice of variables is of paramount interest in the context of the research question. If the idiosyncratic component of the model is overfitted, the systemic risk component can by construction not become significant. Vice versa, if the idiosyncratic part of the model is fitted too generously, it will automatically make the SRM significant. In addressing this trade-off, the work of Demirgüc-Kunt (1989), Wheelock and Wilson (2000), and Cebula (2010) helped shape the model shown in Equation (2.13). Given the scarce occurrence of bank defaults, I choose a probit model with an inverse standard normal link function over the logit model, whose logistic link function has fatter tails, implying higher failure frequencies. $\Phi = \mathbb{P}(D_{i,t} = 1 \mid X = x_{i,t})$ estimates the PD of individual banks and is constrained by the domain $\in \{0, 1\}$.

$$\begin{aligned}\widehat{PD}_{i,t+1} =& \Phi\big(\alpha + \beta_1 CASH_{i,t} + \beta_2 EQT_{i,t} + \beta_3 LOANS_{i,t} + \beta_4 NPL_{i,t} \\ &+ \beta_5 CI_{i,t} + \beta_6 ROID_{i,t} + \beta_7 LTD_{i,t} + \beta_8 ROA_{i,t} + \beta_9 LR_{i,t} + \epsilon_{i,t}\big)\end{aligned} \tag{2.13}$$

The importance of both, capital, and liquidity requirements as under the Basel Accords has been proven time and again (Hugonnier and Morellec (2017)). I thus incorporate cash (CASH) as measure of a bank's liquidity, and Equity (EQT) as a gauge of the loss-absorbing capital. While cash can be used to overcome short-term funding gaps, equity serves as a continuous backstop against losses that can arise from the loan portfolio (LOANS) and are not covered by impairments. As Berger and Bouwman (2013) show, EQT is an especially valuable predictor for the survival of small banks during banking crises. All three measures were logarithmized in order to address the inherent skewness of the observations. The NPL capture the quality of the outstanding loans, while the cost-income-ratio (CI) is a measure of the operating efficiency of the bank. The revenue diversification measure (ROID) approximates how well the bank can sustain a shock to either of its income sources. In line with that the sensitivity to constrained short-term refinancing options is measured as the loan to deposit ratio (LTD). The return on assets (ROA) captures the profitability of the bank, while the balance sheet leverage ratio (LR) addresses the relation of equity to total assets. As argued in Section (2.3) the usage of the regulatory LR would have induced a selection bias to the data set, as it precedes the introduction of Tier 1 capital under the Basel Accords. The aforementioned variables solely grasp the idiosyncratic risk, as shown by Weiß et al. (2014). I thus extend Equation (2.13) individually with the SRM from Section (2.2.2) in order to form Equation (2.14), which holistically captures default risk.

$$\begin{aligned}\widehat{PD}_{i,t+1} =& \Phi\big(\alpha + \beta_1 CASH_{i,t} + \beta_2 EQT_{i,t} + \beta_3 LOANS_{i,t} + \beta_4 NPL_{i,t} \\ & + \beta_5 CI_{i,t} + \beta_6 ROID_{i,t} + \beta_7 LTD_{i,t} + \beta_8 ROA_{i,t} + \beta_9 LR_{i,t} \\ & + \beta_{10} SRM_{i,t} + \epsilon_{i,t}\big)\end{aligned} \tag{2.14}$$

While the incorporation of SRM yields an estimation model that accounts for both, idiosyncratic and systemic risk, it has one remaining shortcoming: the estimated default probability is a metric variable, whereas the default indicator is binary. In order to facilitate comparison between the two, I transform $\widehat{\text{PD}}$ into a dichotomous variable that is one, whenever it exceeds the default boundary λ, respectively zero in all other cases. The cut-off point for λ was chosen at 50 % + ϵ, as to account for the default being more likely than not. In order to address the potential endogeneity problem of reverse causality between high PDs and systemic risk, I use the data of the current quarter t to predict the subsequent quarter $t+1$. The forecasting horizon was chosen in line with the findings of Cole and Gunther (1998) and Breitung and Knüppel (2018), who empirically show that the informativeness of predictions beyond two quarters is questionable at best. In addition, longer lags deplete the model of explanatory power due to the steep default paths, which Vazza and Kraemer (2018) describe.

$$\hat{D}_{i,t+1} = \begin{cases} 1 & \text{for } \Phi\big(\alpha + \beta_1 CASH_{i,t} + \beta_2 EQT_{i,t} + \beta_3 LOANS_{i,t} + \beta_4 NPL_{i,t} \\ & +\beta_5 CI_{i,t} + \beta_6 ROID_{i,t} + \beta_7 LTD_{i,t} + \beta_8 ROA_{i,t} + \beta_9 LR_{i,t} \\ & +\beta_{10} SRM_{i,t} + \epsilon_{i,t}\big) > \lambda \\ 0 & \text{otherwise} \end{cases} \tag{2.15}$$

Doing so allows the calibration of the two-staged testing methodology as discussed before. It is now possible to estimate a bank's PD with a mixed model, which captures idiosyncratic and systemic risk. If the tested SRM are true measures of systemic risk, they should explain the subset of systemically induced bank defaults, and hence increase the explanatory power of Equation (2.15). I thus posit:

Hypothesis 1 *The addition of systemic risk measures improves the predictability of bank failures.*

In order to test this hypothesis, I will evaluate four models in total. A version of Equation (2.13) with only idiosyncratic variables, as well as the amended version from Equation (2.14) which includes ΔCoVaR, MES, and SRISK individually.

The proposed methodology has two major benefits. First, idiosyncratic and systemic risk can build up independently from another. As a result, idiosyncratic risk can be high, while systemic risk is low, and vice versa (recall Figure (2.1)). Because of that, defaults can be triggered by any of the risks individually, such that there is no need for overfitting, as observed in other studies. This characteristic makes for a compelling economic interpretation as a bank can default due to a systemic shock, even when it is sound from a microprudential perspective, or inversely default in a solid macroeconomic environment due to bank specific factors. Furthermore, it works well with the observation of only gradually changing idiosyncratic risk measures, such that the issue of biased standard errors as raised by Mahadeva and Robinson (2004) is of no concern. Second, ex ante assumptions about weighting idiosyncratic and systemic risk become obsolete, as it is implicitly being accounted for by the coefficients of the proposed model. In summary, the postulated methodology describes a parsimonious model, which reduces complexity to a reasonable degree, without becoming incoherent.

2.5 Results

For a more intuitive interpretation of the results, the following coefficients refer to the marginal effects of the probit model. As such, they can be interpreted as the change in the respective PDs. Because they are computed as a derivative, there is no constant to report in the subsequent tables.

Table (2.1) depicts the idiosyncratic base model in line with Equation (2.13) in the first column. It shows that more cash and equity reduce ceteris paribus the probability of bank failure. Given that they constitute additional loss-absorbing capacity, this observation is plausible. At the same time a higher nominal amount of outstanding loans, and a higher percentage of non-performing loans increase, in line with the results of Wheelock and Wilson (2000), the likelihood of bank default in the base model. Lastly, a higher balance sheet leverage ratio, that is more equity relative to total assets, reduces the probability of bank failure. Despite issues such as netting, the topology of balance sheet information appears particularly valuable for estimating the health of a bank. Hence, the findings are in line with Estrella et al. (2000) and Pankoke (2014) who show, that simple risk indicators tend to outperform sophisticated ones.

Table 2.1: Probit regression on default dummy.

	(1) Base Case	(2) MES	(3) ΔCoVaR	(4) SRISK
CASH	-0.0004***	-0.0001*	0.0001	-0.0002**
	(0.0000)	(0.0121)	(0.7424)	(0.0070)
EQT	-0.0008***	-0.0005***	-0.0007	-0.0005***
	(0.0000)	(0.0001)	(0.1393)	(0.0001)
LOANS	0.0011***	0.0006***	0.0010	0.0006***
	(0.0011)	(0.0000)	(0.0975)	(0.0000)
NPL	0.0202***	0.0134***	0.0088*	0.0099***
	(0.0000)	(0.0000)	(0.0295)	(0.0133)
CI	-0.0000	-0.0004***	-0.0000*	-0.0004***
	(0.6592)	(0.0009)	(0.0304)	(0.0009)
ROID	-0.0000	-0.0000*	0.0000	-0.0000
	(0.4966)	(0.5810)	(0.8488)	(0.5762)
LTD	-0.0000	-0.0000	-0.0000	-0.0015
	(0.6065)	(0.6790)	(0.9732)	(0.7005)
ROA	-0.0001	-0.0355***	-0.0007*	-0.0056***
	(0.1725)	(0.0000)	(0.0226)	(0.0000)
LR	-0.0764***	-0.0525***	-0.0283*	-0.0240***
	(0.0000)	(0.0000)	(0.0111)	(0.0000)
MES		-0.0013**		
		(0.0068)		
ΔCoVaR			-0.0039	
			(0.1171)	
SRISK				0.0070**
				(0.0045)
N	1,182,485	850,683	669,346	850,679
BIC	12,702.61	6,839.38	1,739.56	6,864.94

Note: The table above shows the results of the idiosyncratic baseline model in Column (1). It depicts the coefficients, which are used as benchmark for the mixed models that include the SRM MES, ΔCoVaR, and SRISK in Columns (2) to (4). Surprisingly, only MES and SRISK are statistically significant. Economically speaking, only SRISK is a consistent estimator because the sign of MES suggests that higher values coincide with a lower probability of default. Significance is denoted at the 5 %, 1 %, and 0.1 % level.

The adjacent columns two to four depict the results of the mixed models, which incorporate idiosyncratic and systemic risk measures. The findings of the idiosyncratic model from the first column can largely be reinstated. At the same time, CI and ROA become significant. While both, the CI and ROA are statistically significant, the coefficient of ROA is also notably different from zero and lowers the probability of default at an economically relevant level. In the case of MES, and SRISK, the SRM become significant, too. However, their effects are diametrically opposing. Higher SRISK values translate to a higher funding gap under systemic distress, and consequently are associated with a positive sign of the coefficient. At the same time, higher MES values indicate larger negative returns, but contradictorily appear to reduce the likelihood of bank failure. ΔCoVaR fails to generate additional insights how systemic risk transmissions into elevated risk of bank failure. This perception stems from the low significance of the regressors, respectively the absence of significance in the SRM. The results are consistent across all models, as the sign and coefficient of the estimates do not alter. Comparable findings are made by Löffler and Raupach (2013), who also document shortcomings of ΔCoVaR and MES. I proceed to further investigate the predictive power of the models by comparing their forecasts from Equation (2.15) to empirical defaults in Table (2.2).

I initially compute the default frequency (γ) by dividing the number of observed defaults by the total number of observations. Doing so allows me to randomly draw from the population and label banks as defaulted or not with the probability γ, respectively $1 - \gamma$. As a result, I obtain a benchmark scenario to compare the models to. In line with the two-staged identification strategy, I begin the analysis with the base model that only consists of idiosyncratic variables, and then extend it to the mixed models that account for systemic risk as well.

Table 2.2: Accuracy per model.

Random Draw

Empirical	Predictions 0	Predictions 1	Σ
0	1,462,029	1,969	1,463,998
1	1,969	3	1,972
Σ	1,463,998	1,972	1,465,970

Base Case

Empirical	Predictions 0	Predictions 1	Σ
0	1,180,802	201	1,181,003
1	1,388	94	1,482
Σ	1,182,190	296	1,182,485

MES

Empirical	Predictions 0	Predictions 1	Σ
0	849,618	78	849,696
1	936	51	987
Σ	850,554	129	850,683

ΔCoVaR

Empirical	Predictions 0	Predictions 1	Σ
0	669,113	36	669,149
1	175	22	197
Σ	669,288	58	669,346

SRISK

Empirical	Predictions 0	Predictions 1	Σ
0	849,819	74	849,893
1	733	53	786
Σ	850,552	127	850,679

Note: The contingency tables above depict the accuracy of the applied models. Empirical observations are denoted row-wise, whereas the columns contain predictions made from the model. The first table shows the expected results for the full sample. As the number of random draws increases, the likelihood of false positives, respectively false negatives converges towards the squared default frequency (λ). The main diagonal yields the number of correct identifications, of which the sum shall be maximal.

Table (2.2) illustrates the results of this approach in contingency tables, which can be read as square matrices that are henceforth referred to as M. Empirical observations are denoted in the rows (r), whereas predictions can be obtained from the columns (c). Accordingly, the main diagonal (i.e. the elements $m_{1,1}$ and $m_{2,2}$) contains correct predictions and measures the accuracy of the model. From this fact, one can derive an intuitive criterion for ranking the respective models:

$$PS_i = \frac{Tr(M_i)}{N_i} \tag{2.16}$$

The prediction score (PS) of a model i can be computed as the trace (Tr) of the square matrix M standardized by its total number of observations (N). In a perfect model, all observations lie on the main diagonal, such that the trace equals the number of observations. Consequently, the quotient converges towards one for higher predictive powers. Although intuitive, this approach has a central shortcoming by not considering the power, respectively size of the test. I thus extend Equation (2.16) by a subtrahend that considers the remaining elements off the main diagonal to address this concern.

$$PS_i = \underbrace{\frac{Tr(M_i)}{N_i}}_{Accuracy} - \underbrace{\frac{m_{2,1}}{antidiagonal(M_i)}}_{Penalty} \tag{2.17}$$

In the context of the research question, I need to balance the power and size of the test. While false positives may entail significant costs on the bank-level, it is only prudent to predict defaults that do not occur, versus missing them. I address this inherent tension by incorporating a subtrahend that weights the false negatives (i.e. $m_{2,1}$) over all incorrect predictions. The more often a SRM fails to pick up a bank default, the larger $m_{2,1}$, such that the penalizing subtrahend grows, and reduces the prediction score. It becomes obvious, that the score is bound by $\in \{-1, 1\}$, when looking at the extrema. In the first case, where only correct predictions are made, all elements lie on the main diagonal, such that the minuend is one. In the absence of wrong predictions, the antidiagonal converges towards zero, hence yielding a PS of one. In the second case, only false predictions are made, such that the minuend is zero. Assuming the worst case, only false negatives are predicted, such that the subtrahend converges towards one. Given its computation as a difference, the PS becomes minus one. Taken together, I obtain a testable metric that should satisfy two conditions in order to aid the regulatory triage:

Proposition 1 *The prudent model yields better predictions than a random draw (minuend criterion).*

Proposition 2 *The prudent model rather predicts than misses a default (subtrahend criterion).*

Table (2.3) yields the results of computing the PS in the last column. They are obtained by subtracting the penalty term in the third column from the standardized accuracy in the second column. One can see that all models are very accurate as they make plenty of correct predictions in line with the first proposition. However, this observation is unsurprising given that the random draw already classifies 99.73 % of the data correctly. In line with previous deliberations, I thus extend the analysis by the penalty term in the third column. The random draw performs well by construction, which is why I will not further elaborate. ΔCoVaR better discriminates between false positives and negatives than the base case model, which only consists of idiosyncratic variables. Surprisingly, MES and SRISK benefit the accuracy of the predictions, but do not prevent the model from missing defaults. Hence, it appears questionable, that they are generic SRM, which pick up additional defaults from systemic events.

Table 2.3: Computation of the Prediction Score.

	Prediction Score		
	Accuracy	Penalty	PS
Random Draw	0.9973	0.5000	0.4973
Base Case	0.9987	0.8735	0.1252
MES	0.9988	0.9231	0.0757
ΔCoVaR	0.9997	0.8294	0.1703
SRISK	0.9991	0.9083	0.0908

Note: Table (2.3) extents the analysis of Table (2.2) by computing the prediction score (PS) for the respective models as the difference between their accuracy and a penalty for false negatives. Each row corresponds to one model from the analysis. The accuracy in the second column is defined as correct predictions over all predictions. This standardization ensures that it is bound by $\in \{0, 1\}$, where higher values indicate better accuracy. As the third column contains the penalty, higher values indicate worse performance. In detail, it is computed as the false negatives divided by all false predictions from the antidiagonal. The worse the discriminatory power, the higher the penalty, such that bad models converge towards one. The last column contains the PS, which is the difference between the accuracy and penalty. The higher the value, the better, as it indicates a high accuracy and a small penalty at the same time. Vice versa, smaller values indicate worse performing models.

In summary, this section has generated ambiguous evidence for the inclusion of SRM in default prediction models. Although MES and SRISK are statistically significant in explaining bank defaults, they do so inconsistently. While higher values of SRISK are associated with higher default probabilities, the opposite is true for MES. Although ΔCoVaR is not statistically significant, it performs best, when benchmarked against empirical data. All SRM advance the base model with only idiosyncratic variables in the accuracy of the prediction. However, ΔCoVaR is the most prudent in terms of predicting defaults instead of missing them. Given the mixed results of the individual SRM, the pursuing section will further assess the robustness of the results, and explore alternative explanations.

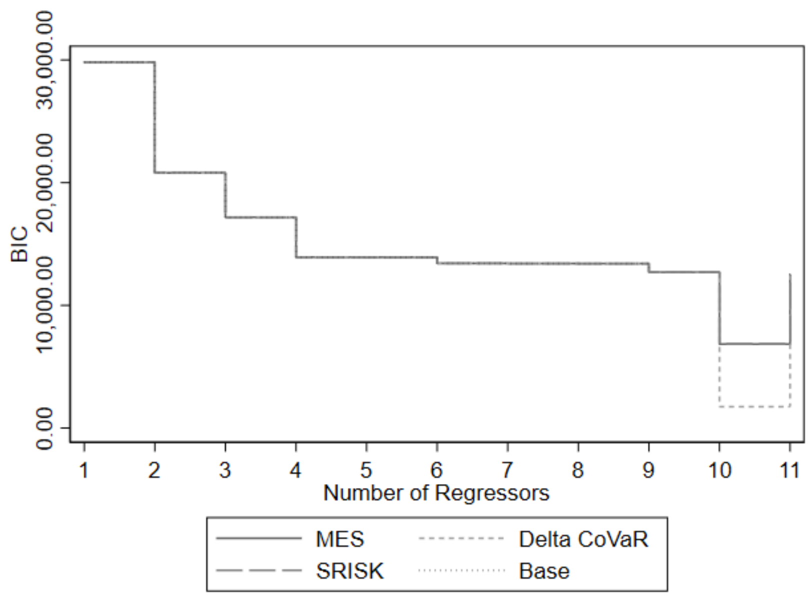
2.6 Robustness

2.6.1 Model Specification

A common critique concerns the model specification. As argued in Section (2.4), there is an inherent trade-off between overfitting the idiosyncratic model, such that systemic risk cannot become significant, and vice versa. To address this problem, the correlation between the regressors was assessed in a first step. In light of an overspecification, excessive correlations between the variables should be observable, as they hint at explaining comparable phenomena with different variables. Table (2.12) in the Appendix depicts the correlations between the used variables. The correlation between equity and loans would be concerning, were it not reflecting structural characteristics of the underlying data set. It consists of many small banks, with low balance sheet leverage ratios. As a result, they show two sides of the same coin. For reasons of controlling, it was decided to use both in the idiosyncratic base model, as omitting either, would constitute an econometrically more severe endogeneity problem. At the other end, the correlation between NPL and ROA is excessively negative. Again, this observation has an economic interpretation to it. ROA is high, when assets are performing, such that interest payments are received, and impairments are deferred. Whenever borrowers become delinquent, no interest payments on the outstanding loan are made, and impairments become mandatory, such that profitability is negatively impacted. The remainder of the variables has encouragingly low correlations, which not only disperses concerns regarding the model specification, but also multicollinearity.

At the other end of the spectrum, underspecification can be a concern when it translates to systemic risk automatically becoming significant. To account for this, the model was amended one variable at a time, considering the Bayesian Information Criterion (BIC) as a performance measure. It was chosen over Akaike's information criterion (AIC), as the latter does not penalize for the number of observations, which varies for the respective models due to data constraints. Because the model accuracy improves with every additional variable that possesses information value, a convex function is expected, where the BIC decreases, until inefficient regressors are added. Hence, a correctly specified model should lie at the minimum of such function. The plot of this iterative addition of variables is depicted in Figure (2.3) below. It has two favorable take aways in line with the argumentation of this paper. First, the addition of the SRM significantly improves the model performance, as visualized by the strong decline in BIC with the addition of a tenth regressor. Second, incorporating further regressors does not benefit the model performance and confirms the convexity, which can be understood as evidence against an omitted variable bias. The numbers on the abscissa correspond to the variables in Equation (2.14).

Figure 2.3: Model performance after adding additional regressors.



2.6.2 Size Effects

Given the multitude of different types of banks, another issue of the applied model could be unaccounted for heterogeneity. It might manifest in systematic differences in the causes of default for small and large banks. Duffie (2011) corroborates this concern, by documenting a dedicated default channel exclusively for larger clearing banks. To investigate this possibility, the sample was divided in four peer groups based on their size as measured by total assets. The rational behind it is that four peer groups would allow to cluster the banks in sufficient granularity above, respectively below the median bank size, while not creating meaningless subsets without bank failures. I generate both, numerical and graphical evidence for this conjecture in the following.

Figure 2.4: Number of defaults per peer group over time.

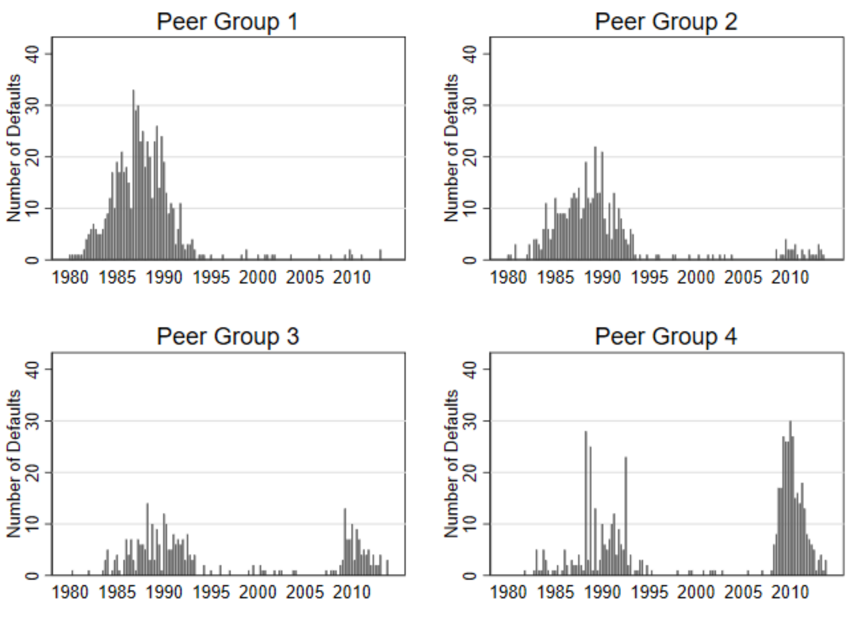


Figure (2.4) reveals noteworthy differences between the bank types. The number of defaults per year for the lower quartile of banks is depicted by peer group one in the top left graph, whereas the largest quartile is illustrated by peer group four in the bottom right graph. As can be seen, default clustering occurred at different times for different types of banks. In line with the results of Berger and Bouwman (2013), smaller banks are sensitive toward credit risk shocks, such as the S&L Crisis, while larger banks appear to be stronger influenced by market risk shocks, such as the Subprime Crisis. Consequently, default prediction with established models is pyrrhic. The explanatory power of every variable is diminished by the fact that it has to account for both, sound and ailing banks alike, and hence gives further credibility to the incorporation of SRM.

This observation can be reinstated using the G-test, which is an extension of the χ^2-test. It works especially well in large data sets (i.e. $\geq$ 1,000 observations), as it is the case here. By comparing the observed defaults in each peer group to the theoretical observations under the null hypothesis that the defaults are uniformly distributed, I find evidence of excess defaults in peer group one, respectively a deficit in peer group three. This difference is testable under the G-test, which is computed as shown in Equation (2.18):

$$G = 2\sum_{i=1}^{N} O_i \times ln\Big(\frac{O_i}{E_i}\Big) \tag{2.18}$$

In it O_i denotes the number of empirical observations in peer group i, while E_i represents the number of expected observations. Together with the degrees of freedom, one can compute the test statistic using a χ^2-distribution. Given the four peer groups, which are divided by the two categories (i.e. default and non default), one can formally arrange the obtained matrix M in the form $\mathrm{M} \in \mathbb{R}^{r \times c}$, where r denotes the number of rows, and c the number of columns. Doing so allows me to compute the degrees of freedom as $(r-1) \times (c-1)$. From this information I infer at the 99.9 % confidence level that bank defaults are not uniformly distributed.

Table 2.4: G-Test for peer groups.

Panel A:

	Observed Values (O_i)	
	Non Default	Default
Peer Group 1	295,465.00	647.00
Peer Group 2	365,832.00	448.00
Peer Group 3	395,620.00	321.00
Peer Group 4	406,818.00	541.00

Panel B:

	Expected Values (E_i)	
	Non Default	Default
Peer Group 1	295,716.63	395.37
Peer Group 2	365,790.94	489.06
Peer Group 3	395,412.34	528.66
Peer Group 4	406,815.09	543.91

Note: The table above shows the observed number of (non) defaults for each of the four respective peer groups in Panel A. Panel B contains the expected frequency of (non) defaults, under the assumption that their occurrence is uniformly distributed between the peer groups. By contrasting the observed and expected frequencies using the G-test, I can compute a test statistic of the following form as in Equation (2.18). I find that the defaults are not uniformly distributed between the peer groups at the 99.9 % confidence level, which suggests the presence of substantial heterogeneity between the respective subsets.

A noteworthy observation can be made for peer group three, which is the least impacted in terms of default frequency. While it is susceptible to both, the S&L Crisis of the 1980s, as well as the Subprime Crisis of 2008, the absolute number of defaults is the lowest for all peer groups. This finding suggests that bank size has crucial implications for their financial stability. On closer inspection, it is not the bank size, but the business model, that leads to it, which determines the banks' exposure to the respective crises. Smaller banks were vulnerable to credit risk, which materialized in numerous defaults during the S&L Crisis. Larger banks at the other end of the spectrum were only marginally impacted from this risk for two reasons. First, they can better diversify their loan portfolio, such that they can fully diversify the idiosyncratic loan risk in theory, as argued by Calomiris and Mason (2003). Second, the practical limitations in the form of a loan portfolio that tends towards infinite granularity can be circumvented by securitizing exposures that are difficult to hedge. At the same time, this transformation of credit risk exposed them to the entailed market risk. When it materialized during the Subprime Crisis of 2008, the default frequency increased, as illustrated in the bottom right plot of Figure (2.4). The diversification in risk dimensions (i.e. credit and market risk) as embodied by peer group three thus suggests that it yields substantial explanatory power in terms of financial stability.

Köhler (2015) attests to this theory by documenting that smaller banks become more resilient, when they increase their marginal non-interest income, and hence develop towards diversified banks, such as those in peer group three. Wagner (2010) adds in this context, that diversification can have drawbacks if banks become too similar, such that all of them are simultaneously exposed to the same shocks.

Furthermore, Figure (2.4) yields graphic evidence in favor of the "regulation hypothesis" of Weiß et al. (2014). They argue that banks, which contributed strongly during the last financial crisis, are being regulated more efficiently ex post, such that their contribution to the next crisis is disproportionally low. In particular, peer groups one and two were more strongly regulated after the S&L Crisis, such that they only marginally added to the 2008 financial crisis. I further validate my deliberations by amending the regressions from Equation (2.14) with a dummy for each of the peer groups. In light of the theorized explanation, the dummies should be significant in the regression.

Table 2.5: Probit regression on the default dummy.

	(1) MES	(2) ΔCoVaR	(3) SRISK
CASH	-0.0002**	0.0000	-0.0002**
	(0.0042)	(0.7647)	(0.0029)
EQT	-0.0005***	-0.0007	-0.0005***
	(0.0001)	(0.1866)	(0.0000)
LOANS	0.0006***	0.0008	0.0006***
	(0.0000)	(0.1457)	(0.0000)
NPL	0.0132***	0.0008	0.0131***
	(0.0000)	(0.0555)	(0.0000)
CI	-0.0000***	-0.0000	-0.0000***
	(0.0005)	(0.2738)	(0.0005)
ROID	-0.0000	-0.0000	-0.0000
	(0.6139)	(0.2520)	(0.6122)
LTD	-0.0000	-0.0000	-0.0000
	(0.5327)	(0.9801)	(0.5179)
ROA	-0.0352***	-0.0006*	-0.0363***
	(0.0000)	(0.0341)	(0.0000)
LR	-0.0519***	-0.0254*	-0.0522***
	(0.0000)	(0.0200)	(0.0000)
MES	-0.0014**		
	(0.0064)		
ΔCoVaR		-0.0005	
		(0.7930)	
SRISK			0.0068**
			(0.0053)
Peergroup (2)	-0.0006**	-0.0001	-0.0005**
	(0.0013)	(0.6875)	(0.0021)
Peergroup (3)	-0.0005**	-0.0005	-0.0004*
	(0.0067)	(0.2249)	(0.0139)
Peergroup (4)	-0.00003	-0.0006	-0.0002
	(0.2812)	(0.1908)	(0.4902)
N	850,683	669,346	850,679
BIC	6,632.64	1,788.98	6,640.01

Note: The table above shows the results of the idiosyncratic and mixed models with dummies for the respective peer groups. Against the common too-big-to-fail-discussion, it suggests that smaller banks are less likely to fail, instead of large banks, which are implicitly insured by the government. Significance is denoted at the 5 %, 1 %, and 0.1 % level.

As can be inferred from Table (2.5), the added peer group dummies are only significant for the smaller banks, with fading statistical and economic significance towards the largest banks. While this observation is in line with the discussed "regulation hypothesis" of Weiß et al. (2014), it generates evidence against the often discussed too-big-to-fail-problematic. A closer inspection shows that the explanatory power of the model as measured by the BIC deteriorates for MES and SRISK, when incorporating the peer group dummies. One can thus infer in line with Figure (2.3), that the previous models from Equations (2.13) and (2.14) already contain the necessary variables to explain the observations. The absence of changes in the sign or magnitude of the coefficients attests to this interpretation. In the interest of completeness, it shall be said that the ΔCoVaR model improves, as EQT and LOANS become significant. However, this change only puts the model on a level playing field with MES and SRISK. It might thus relate to a multiple comparison bias, where the significance is induced by the margin of error of the respective confidence levels.

I further investigate the effect of bank size and the resulting business model, by re-estimating Equations (2.13) and (2.14) for the subsets of the individual peer groups in Tables (2.6) to (2.8). If the postulated model identifies the true transmission channels of bank default, the findings should be replicable on a peer group level, irrespective of previous deliberations.

Table 2.6: Regression with MES by quartiles.

	(1)	(2)	(3)	(4)	(5)
	Quartiles				
	1st	2nd	3rd	4th	All
CASH	-0.0002	-0.0003*	-0.0002*	-0.0002	-0.0001*
	(0.4105)	(0.0144)	(0.0059)	(0.1006)	(0.0121)
EQT	-0.0021***	-0.0004	-0.0001	-0.0005	-0.0005***
	(0.0002)	(0.0522)	(0.3885)	(0.6394)	(0.0001)
LOANS	0.0021**	0.0011***	0.0004	0.0008	0.0006***
	(0.0028)	(0.0000)	(0.0889)	(0.5668)	(0.0000)
NPL	0.0284***	0.0105***	0.0076***	0.0163	0.0134***
	(0.0000)	(0.0000)	(0.0000)	(0.0568)	(0.0000)
CI	0.0001	-0.0000*	-0.0004***	-0.0004***	-0.0004***
	(0.0679)	(0.0125)	(0.0001)	(0.0005)	(0.0009)
ROID	0.0001	0.0002	-0.0000	-0.0000	-0.0000
	(0.8465)	(0.6818)	(0.8392)	(0.9177)	(0.5810)
LTD	0.0000***	0.0000*	-0.0004	-0.0008	-0.0000
	(0.0005)	(0.0324)	(0.4772)	(0.6905)	(0.6790)
ROA	-0.0391	-0.0154	-0.0313***	-0.0350**	-0.0355***
	(0.1489)	(0.1019)	(0.0000)	(0.0016)	(0.0000)
LR	-0.0843***	-0.0349**	-0.0359***	-0.0757***	-0.0525***
	(0.0005)	(0.0013)	(0.0000)	(0.0000)	(0.0000)
MES	-0.0025	-0.0010	-0.0003	-0.0031	-0.0013**
	(0.2262)	(0.1616)	(0.7023)	(0.0883)	(0.0068)
N	114,707	199,733	254,546	281,697	850,683
BIC	1,623.13	1,261.55	1,491.90	2,701.95	6,839.38

Note: The table above shows the results of Equation (2.14) estimated for subsamples consisting of the respective quartiles of the population. I find that the SRM is only significant for the whole population, but not the subsamples. Given that the signed value of MES is inconsistent, as it points in the wrong direction, this observation is unsurprising and attests to the shortcomings of MES. At the same time, NPL and LR are throughout the quartiles highly significant, and appear to have high explanatory power of bank failure. Significance is denoted at the 5 %, 1 %, and 0.1 % level.

Table 2.7: Regression with ΔCoVaR by quartiles.

	(1)	(2)	(3)	(4)	(5)
	Quartiles				
	1st	2nd	3rd	4th	All
CASH	0.0001	-0.0000	0.0001	0.0001	0.0001
	(0.8342)	(0.7430)	(0.3419)	(0.5464)	(0.7424)
EQT	-0.0002	-0.0001	-0.0004	-0.0014*	-0.0007
	(0.8127)	(0.1285)	(0.1545)	(0.0176)	(0.1393)
LOANS	0.0004	0.0006*	0.0006	0.0016**	0.0010
	(0.8149)	(0.0201)	(0.0533)	(0.0049)	(0.0975)
NPL	0.0012	0.0026**	0.0048**	0.0124**	0.0088*
	(0.7650)	(0.0068)	(0.0049)	(0.0000)	(0.0295)
CI	-0.0000	-0.0000*	-0.0002**	-0.0003***	-0.000*
	(0.7967)	(0.0177)	(0.0084)	(0.0000)	(0.0304)
ROID	-0.0001	0.0002	-0.0000*	-0.0000	0.0000
	(0.7763)	(0.4777)	(0.0272)	(0.5891)	(0.8488)
LTD	-0.0000	0.0000	0.0000	-0.0000	-0.0000
	(0.9996)	(0.9457)	(0.1067)	(0.5685)	(0.9732)
ROA	-0.0000	-0.0009**	-0.0031**	-0.0152	-0.0007*
	(0.7256)	(0.0095)	(0.0018)	(0.2049)	(0.0026)
LR	0.0009	-0.0079	-0.0514*	-0.0492**	-0.0283*
	(0.8913)	(0.0622)	(0.0260)	(0.0018)	(0.0111)
ΔCoVaR	-0.0014	0.0005	-0.0012	-0.0068***	-0.0039
	(0.8434)	(0.6658)	(0.2575)	(0.0000)	(0.1171)
N	89,831	162,218	200,606	216,691	669,346
BIC	199.92	327.17	494.25	974.37	1,739.56

Note: The table above shows the results of Equation (2.14) estimated for subsamples consisting of the respective quartiles of the population. The variable of interest, ΔCoVaR is insignificant throughout all, but the last quartile. One can observe a trend moving from the smallest to the largest banks, showing that ΔCoVaR becomes continuously more significant for them. This finding suggests that the measure might work well for larger banks, but fails to grasp the systemic risk of smaller banks. Significance is denoted at the 5 %, 1 %, and 0.1 % level.

Table 2.8: Regression with SRISK by quartiles.

	(1)	(2)	(3)	(4)	(5)
	Quartiles				
	1^{st}	2^{nd}	3^{rd}	4^{th}	All
CASH	-0.0002	-0.0003*	-0.0003**	-0.0002	-0.0002**
	(0.3317)	(0.0140)	(0.0030)	(0.7061)	(0.0070)
EQT	-0.0021***	-0.0004	-0.0004	-0.0006	-0.0005***
	(0.0002)	(0.0507)	(0.4227)	(0.7652)	(0.0001)
LOANS	0.0017*	0.0011***	0.0003	0.0008	0.0006***
	(0.0156)	(0.0000)	(0.1659)	(0.7941)	(0.0000)
NPL	0.0273***	0.0099***	0.0074***	0.0169	0.0099***
	(0.0000)	(0.0000)	(0.0000)	(0.6689)	(0.0133)
CI	0.0001	-0.0000*	-0.0004***	-0.0005	-0.0004***
	(0.0681)	(0.0127)	(0.0000)	(0.6484)	(0.0009)
ROID	0.0002	0.0002	0.0000	0.0000	-0.0000
	(0.7905)	(0.6924)	(0.9442)	(0.9784)	(0.5762)
LTD	0.0000***	0.0000*	-0.0004	-0.0008	-0.0015
	(0.0003)	(0.0257)	(0.5278)	(0.8677)	(0.7005)
ROA	-0.0342	-0.0157	-0.0305***	-0.0380***	-0.0056***
	(0.1985)	(0.0967)	(0.0000)	(0.0009)	(0.0000)
LR	-0.0802***	-0.0352**	-0.0360***	-0.0794	-0.0240***
	(0.0009)	(0.0017)	(0.0000)	(0.6079)	(0.0000)
SRISK	34.8195**	0.8236***	0.8815*	0.0080	0.0070**
	(0.0090)	(0.0002)	(0.0143)	(0.6780)	(0.0045)
N	114,704	199,733	254,545	281,697	850,679
BIC	1,600.12	1,261.79	1,465.14	2,732.52	6,864.94

Note: The table above shows the results of Equation (2.14) estimated for subsamples consisting of the respective quartiles of the population. I find that SRISK is highly significant for all but the largest banks. The coefficient for the smallest banks stands out from all others and is thus of particular interest. An explanation for its size relates to an economic interpretation of SRISK. It measures the funding gap of a bank, which banks can only narrow through retaining earnings or issuing new capital. Larger banks may close this funding gap more efficiently, explaining the declining significance towards the largest quartile. Significance is denoted at the 5 %, 1 %, and 0.1 % level.

As indicated by earlier results, ΔCoVaR fails to enhance our understanding of bank defaults, which is why I will focus on Tables (2.6) and (2.8) in the interest of conciseness. A notable observation concerns the revenue diversification measure ROID, which is never significant. At the same time NPL is always significant. It thus appears as if the distribution of revenues is negligible in the context of bank stability, whereas the opposite is true for the distribution of risks (i.e. credit and market risk). In line with this interpretation, LOANS is significant throughout all models for the smallest banks, which suggests that the exposure to credit risk has noteworthy explanatory power. Likewise, LTD is significant for banks below the median, which suggests that stable funding contributes to their resilience. The SRM, which are the focal point of this work, unveil interesting insights on further inspection. MES is for neither of the peer groups significant, while SRISK is significant for all but the largest. This observation further attests to the advantages of SRISK. It is not only a consistent measure, as the sign is in line with economic theory, but also possesses high explanatory power as illustrated by the accuracy ratio, as well as the lower respective BIC. Lastly, SRISK offers an intuitive interpretation in light of its significance. The coefficient for the smallest banks is distinguishably large from other coefficients. However, the explanation lies at hand: SRISK measures the funding gap between a banks debt and its equity. Banks can only close this gap by either issuing additional capital, or retaining earnings. This limitation makes it difficult for them to overcome their funding deficit, and hence explains the high explanatory power of SRISK on default. At the same time, it appears that larger banks can close this funding gap more efficiently, which attests to the growing insignificance for bigger institutions. There are multiple possible explanation for this observation. One relates to larger banks being less opaque, because they are subject to stricter disclosure regimes, (e.g. the Dodd-Frank Act or Pillar III requirements of the Basel Accords) and disciplined by the capital market (see Demirgüç-Kunt et al. (2008)). As a result, an agreeable valuation for closing the funding gap can be obtained more easily. Another explanation relates to implicit state guarantees, which make funding more accessible for larger "too-big-to-fail" banks (Tsesmelidakis and Merton (2013)).

2.6.3 Sample Variation

In order to account for the unbalanced nature of the sample, I conduct two further robustness tests, which reduce, respectively amend the number of defaults. I begin by truncating the independent variables in line with the interquartile range (IQR) as suggested by Tukey (1977) in order to assess the influence of outliers on the results. Winsorization would not yield meaningful values in this instance, as it replaces the outliers with observations at a given percentile. As such, it does not have the potential to generate new insights given the size of this data set, which dilutes the impact of individual outliers in the first place. Truncation, however, omits the outliers by taking the IQR into consideration, which better grasps the inherent skewness of a data set with both, large and small banks alike. As such the body of the distribution becomes more pronounced at the expense of the tails, which should contain the most, respectively least resilient banks. Hence, this robustness test attains an economic interpretation, as it emphasizes the effect on the average bank. Given the binary nature of the dependent variable, it was not changed.

Table 2.9: Regression on the default dummy for the full and truncated sample.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Base Model		MES		ΔCoVaR		SRISK	
	Full	Truncated	Full	Truncated	Full	Truncated	Full	Truncated
CASH	-0.0004***	-0.0001**	-0.0001*	-0.0000	0.0001	-0.0000	-0.0002***	-0.0000
	(0.0000)	(0.0015)	(0.0121)	(0.0652)	(0.7424)	(0.8596)	(0.0070)	(0.0625)
EQT	-0.0008***	-0.0001	-0.0005***	-0.0001	-0.0007	-0.0000	-0.0005***	-0.0001
	(0.0000)	(0.5543)	(0.0001)	(0.3472)	(0.1393)	(0.8934)	(0.0001	(0.3384)
LOANS	0.0011***	0.0001	0.0006***	0.0001	0.0010	0.0000	0.0006***	0.0001
	(0.0011)	(0.3444)	(0.0000)	(0.9938)	(0.0975)	(0.7563)	(0.0000)	(0.2197)
NPL	0.0202***	0.0022*	0.0134***	0.0011	0.0088*	-0.0001	0.0099***	0.0011
	(0.0000)	(0.0442)	(0.0000)	(0.0673)	(0.0295)	(0.7611)	(0.0133)	(0.0682)
CI	-0.0000	-0.0000	-0.0004***	-0.0000	-0.0000*	0.0000	-0.0004***	-0.0000
	(0.6592)	(0.8289)	(0.0009)	(0.8297)	(0.0304)	(0.9708)	(0.0009)	(0.8294)
ROID	-0.0000	0.0005*	-0.0000*	0.0001	0.0000	0.0000	-0.0000	0.0001
	(0.4966)	(0.0154)	(0.5810)	(0.5241)	(0.8488)	(0.6445)	(0.5762)	(0.5245)
LTD	-0.0000	-0.0000	-0.0000	-0.0004*	-0.0000	-0.0000	-0.0015	-0.0004
	(0.6065)	(0.8841)	(0.6790)	(0.0168)	(0.9732)	(0.9008)	(0.7005)	(0.5762)
ROA	-0.0001	-0.0000	-0.0355***	-0.0054**	-0.0007*	-0.0001	-0.0056***	-0.0054**
	(0.1725)	(0.9215)	(0.0000)	(0.0028)	(0.0226)	(0.8093)	(0.0000)	(0.0028)
LR	-0.0764***	-0.0154***	-0.0525***	-0.0054**	-0.0283*	-0.0008	-0.0240***	-0.0055**
	(0.0000)	(0.0000)	(0.0000)	(0.0011)	(0.0111)	(0.2233)	(0.0000)	(0.0011)
MES			-0.0013**	-0.0000				
			(0.0068)	(0.9764)				
ΔCoVaR					-0.0039	-0.0001		
					(0.1171)	(0.4660)		
SRISK							0.0070**	-0.0099
							(0.0045)	(0.8823)
N	1,182,485	897,406	850,683	664,908	669,346	522,639	850,679	664,908
BIC	12,702.61	1,657.75	6,839.38	1,080.20	1,739.56	294.27	6,864.94	1,080.18

Note: The table above shows the results of Equation (2.14) estimated for the full sample in the odd-numbered columns, whereas the results of the truncated population are shown in the even-numbered columns. It is obvious to the eye, that the significance of most variables has vanished. The explanation for this observation lays at hand: the banks most likely to fail have been dropped from the sample due to the truncation technique. As a result, only resilient banks are used to re-estimate the model. However, as Cole and Gunther (1995) show, indicators of bank distress, are not inversely indicators of bank stability. Thus, their findings explain the strong reduction in significance of the explanatory variables. At the same time, the BIC has declined dramatically, which is logical because the model now only has to fit sound banks. Significance is denoted at the 5 %, 1 %, and 0.1 % level.

To ease legibility, Table (2.9) depicts the results from the full sample in the odd-numbered columns, whereas the results of the truncated model are shown in the adjacent even-numbered columns. Comparing the respective models yields striking insights: (i) all of the tested SRM become insignificant, while (ii) most of the idiosyncratic variables have lost their explanatory power, too. This observation would be troubling, had it not a natural explanation. The truncation has depleted the model of its most extreme values, which were also the most failure prone. As a result, the number of observed defaults has gone down from 2,044 to 124. The lack of explanatory power is thus in favor of the postulated hypothesis because the estimated model does not describe characteristics of failing banks, but of sound banks. The strong decline in BIC attests to this interpretation, as the model does not longer have to discriminate between sound and ailing banks, but instead focuses on identifying characteristics of resilient banks almost exclusively. In line with this explanation, Cole and Gunther (1995) show that the variables that explain bank failure cannot be interpreted bidirectionally because they do not predict bank survival.

While the previous robustness test has reduced the number of observed defaults, the inclusion of “quasi defaults” is investigated as a further mean of confirming the results. The Troubled Asset Relief Program (TARP) constitutes a unique opportunity of amending the sample population with additional defaults. I increase the number of observed bankruptcies by counting banks, which were subject to TARP, as failed. Consequently, the data set becomes more balanced because there are more defaults, but also less non defaults, since observations after these theoretical defaults have been dropped.

Table 2.10: Regression on the default dummy for the full and amended sample.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Base Model		MES		ΔCoVaR		SRISK	
	Full	TARP	Full	TARP	Full	TARP	Full	TARP
CASH	-0.0004***	-0.0003***	-0.0001*	-0.0001	0.0001	0.0001	-0.0002***	-0.0001
	(0.0000)	(0.0000)	(0.0121)	(0.0870)	(0.7424)	(0.1989)	(0.0070)	(0.0791)
EQT	-0.0008***	-0.0015***	-0.0005***	-0.0007***	-0.0007	-0.0011***	-0.0005***	-0.0007***
	(0.0000)	(0.0000)	(0.0001)	(0.0000)	(0.1393)	(0.0009)	(0.0001	(0.0000)
LOANS	0.0011***	0.0018***	0.0006***	0.0008***	0.0010	0.0014***	0.0006***	0.0008***
	(0.0011)	(0.0000)	(0.0000)	(0.0000)	(0.0975)	(0.0002)	(0.0000)	(0.0000)
NPL	0.0202***	0.0238***	0.0134***	0.0138***	0.0088*	0.0104***	0.0099***	0.0137***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0295)	(0.0000)	(0.0133)	(0.0000)
CI	-0.0000	-0.0000	-0.0004***	-0.0000***	-0.0000*	-0.0000**	-0.0004***	-0.0000***
	(0.6592)	(0.1176)	(0.0009)	(0.0008)	(0.0304)	(0.0023)	(0.0009)	(0.0008)
ROID	-0.0000	-0.0000	-0.0000*	-0.0000	0.0000	-0.0000	-0.0000	-0.0000
	(0.4966)	(0.2850)	(0.5810)	(0.6451)	(0.8488)	(0.3465)	(0.5762)	(0.6424)
LTD	-0.0000	-0.0000*	-0.0000	-0.0000	-0.0000	-0.0000	-0.0015	-0.0000
	(0.6065)	(0.0339)	(0.6790)	(0.6408)	(0.9732)	(0.5348)	(0.7005)	(0.6152)
ROA	-0.0001	-0.0001***	-0.0355***	-0.0385***	-0.0007*	-0.0006***	-0.0056***	-0.0388***
	(0.1725)	(0.0000)	(0.0000)	(0.0000)	(0.0226)	(0.0000)	(0.0000)	(0.0000)
LR	-0.0764***	-0.0053***	-0.0525***	-0.0415***	-0.0283*	-0.0167**	-0.0240***	-0.0417***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0111)	(0.0035)	(0.0000)	(0.0000)
MES			-0.0013**	-0.0005				
			(0.0068)	(0.1895)				
ΔCoVaR					-0.0039	-0.0005		
					(0.1171)	(0.6873)		
SRISK							0.0070**	-0.0004
							(0.0045)	(0.9204)
N	1,182,485	1,181,238	850,683	835,404	669,346	663,016	850,679	835,404
BIC	12,702.61	14,164.74	6,839.38	7,054.37	1,739.56	2,086.33	6,864.94	7,059.75

Note: The table above shows the results of Equation (2.14) estimated for the full sample in the odd-numbered columns, whereas the results including "quasi defaults" are shown in the even-numbered columns. One can observe that the SRM become insignificant for the new population. Unlike the instance of the truncated sample, the idiosyncratic measures continue to constitute explanatory power. An explanation for this observation might stem from a possible self-selection issue. TARP was designed to aid banks with toxic assets, and as such the inclusion of "quasi defaults" has amended the data set with many banks that were unstable from an idiosyncratic perspective. As a result, these banks opted to participate in TARP and hence bias the new model, by calibrating it on banks, that were exclusively failure prone for bank specific, instead of systemic reasons. It is worth noting that the overall number of observations in the TARP sample declines, because more defaults occur, such that there are less non-default observations. Significance is denoted at the 5 %, 1 %, and 0.1 % level.

Table (2.10) presents the results of these “quasi defaults”. It follows the same structure as Table (2.9) and shows the previous results in the odd-numbered columns, while the adjacent even-numbered columns yield the coefficients of the TARP model. Again, the tested SRM become insignificant, when accounting for “quasi defaults”. At the same time, the idiosyncratic measures remain significant with similar coefficients and static signs. A possible explanation for this observation relates to the TARP program itself. It allowed banks with overburdening ratios of illiquid or non-performing assets to unload them at the U.S. Treasury. Thus, the TARP program describes a self-selection mechanism, where banks that deemed themselves at the center of idiosyncratic risk through their loan portfolios remedied this risk through loan sales.

Another mean of generating additional defaults for back-testing the model is described by Cole and White (2012). They generate “technical defaults”, whenever the sum of a bank’s equity and loan loss reserves does not cover at least half the non-performing assets. Doing so identifies another 708 banks as defaulted in this data set. However, on closer inspection, I find that “technical defaults” do not identify additional bank failures, but rather reveal them up to three quarters in advance. This observation is confirmed by re-estimating the model. In untabulated results, I find the coefficients to be similar to those from Table (2.1), except for MES, which also becomes insignificant.

Taken together, this section attest to the robustness of the presented results, as neither reducing, nor amending the number of observed defaults generates contradicting evidence.

2.7 Conclusion

Bank defaults continue to cause severe economic distortions. Hence, identifying troubled banks and taking precautions is in the interest of all stakeholders of the economy. This paper goes beyond the current literature by extending the understanding of bank failure. In the status quo, bank defaults are mostly regarded as the culmination of idiosyncratic problems on the bank-level. However, this understanding omits the role of systemic risk, despite bank defaults clustering around systemic events such as the S&L Crisis or the Subprime Crisis. I remedy this shortcoming by testing the influence of SRM in the context of default prediction.

In a first step, a default prediction model, which only consists of idiosyncratic risk measures is derived. If systemic risk triggers bankruptcy in financial institutions, this base model should fail to account for such defaults. Hence, I posit the hypothesis that if the established SRM are true measures of systemic distress, their incorporation should increase the accuracy of default prediction models.

The generated results show that ΔCoVaR is insignificant in the subsequent econometric models, but performs well, when compared to de facto defaults. While MES and SRISK are both statistically significant, only SRISK consistently improves the accuracy of default prediction. The negative sign of the MES coefficient inconsistently suggests that more systemic risk makes a bank less failure prone. Thus substantive evidence against the research question is generated for all SRM, but SRISK.

In order to assess the validity of the results three strands of robustness tests are conducted. First, the model specification is investigated. If the idiosyncratic model were overly specified, the SRM could by construction not become significant. I generate evidence against this possibility by investigating the underlying correlations and demonstrating the convexity of the applied models in terms of their BIC coefficient. Second, I test for unaccounted heterogeneity from bank size. While there is evidence that size determines the exposure of banks to different types of financial crises, the postulated models function irrespective of bank size. The results again attest to the quality of SRISK as prevailing measure of systemic risk. Third, the influence of the chosen sample is analyzed. Neither reducing, nor amending the number of observed defaults in the data set challenges the previous results. Hence, the findings suggest that the selected variables are true predictors of bank distress, be it in the form of idiosyncratic or systemic risk. From this observation, it becomes evident, that a single indicator should never be the center of predicting bank failure. Instead, a multitude of indicators has to be assessed, where the literature suggests that simpler indicators perform better.

The results of this paper are intriguing, because they entail profound policy implications. They challenge the status quo of measuring systemic risk by pointing out the shortcomings of the established measures. In doing so, the regulatory triage is aided by suggesting that if any of the measures is to be used, it should be SRISK. Another pivotal insight of this work relates to LR and ROA as indicators of bank distress. Both variables have proven robust throughout the conducted analyses and hence present themselves as credible indicators of financial distress in a broader context.

As the name suggests, systemic risk stems from the system itself. It thus appears prudent for future research to reinstate the made findings in bank-based systems, instead of a market-based system, as it is the case here. Doing so might reveal other interesting transmission mechanisms of default risk in banks. The European sovereign debt crisis emerges as a predestine opportunity for doing so, given that many European banks operate in bank-based economies. Another interesting line of thought opens up around the results of the subsampling. Apparently, there is a sweet spot in terms of diversification between credit and market risk, where bank resilience is maximal. Narrowing down this corridor of bank size and describing the underlying mechanics in detail is likely to yield important insights. Lastly, “technical defaults” appear to hint at future bank failures, such that it might be worth revisiting them.

2.8 Appendix

Table 2.11: Descriptive statistics.

	Min	Q_{25}	Median	Mean	Q_{75}	Max	SD
Multiple	0.5785	0.9538	1.1668	1.2069	1.4456	2.2110	0.3776
ln(Cash)	5.4116	7.2772	8.0662	8.2202	8.9660	12.9557	1.4627
ln(Equity)	6.1225	7.8087	8.6078	8.7488	9.4989	13.1140	1.3951
ln(Loans)	7.4061	9.5141	10.3879	10.5256	11.3829	15.1015	1.5365
NPL	0.0000	0.0033	0.0101	0.0190	0.0236	0.1309	0.0284
CI	0.5158	0.7446	0.8127	0.8614	0.8822	1.7237	5.6694
ROID	0.0000	0.0966	0.1563	0.1754	0.2412	0.7694	4.3133
LTD	0.1734	0.5474	0.6764	0.7779	0.7968	1.2653	2.0072
ROA	-0.0133	0.0024	0.0039	0.0035	0.0052	0.0146	1.0472
LR	0.0378	0.0755	0.0895	0.1026	0.1102	0.3656	0.0706
MES	-0.3885	-0.2388	-0.1996	-0.1974	-0.1810	-0.0342	0.3338
ΔCoVaR	-0.1644	-0.0956	-0.0719	-0.0696	-0.0405	0.0002	0.0389
SRISK	-141,688.0938	-4,120.0486	-1,395.8403	-17,888.1421	-295.0422	28,260.0742	586,267.9346

Note: The table above depicts the descriptive statistics of the variables used in the probit models. Cash, equity, and loans were logarithmized in order to account for the inherent skewness. Standardized values of loans and equity can be obtained from LTD and LR. The values of SRISK are significantly larger than the remaining variables, which is unproblematic due to the absence of linear relationships in the probit model. Furthermore, standardizing them would have depleted the SRISK measure of banks, which do not have a capital shortfall. Hence, it was decided against doing so, in order to keep more explanatory variance in this variable.

Table 2.12: Correlation table.

	idiosyncratic base model										systemic component		
	Default	CASH	EQT	LOANS	NPL	CI	ROID	LTD	ROA	LR	MES	ΔCoVaR	SRISK
Default	1.0000												
CASH	0.0231	1.0000											
EQT	-0.0017	0.8449	1.0000										
LOANS	0.0183	0.8324	0.9331	1.0000									
NPL	0.1018	0.0585	-0.0092	-0.0077	1.0000								
CI	-0.0101	0.0317	0.0644	0.0439	-0.0362	1.0000							
ROID	-0.0003	0.0040	0.0110	-0.0195	-0.0008	0.0022	1.0000						
LTD	-0.0001	0.0167	0.0219	0.0234	0.0016	0.0190	-0.0001	1.0000					
ROA	-0.1100	-0.0225	0.0595	0.0133	-0.2445	0.1117	0.0806	-0.0023	1.0000				
LR	-0.0334	-0.1249	0.0906	-0.1815	-0.0090	0.0952	0.0977	-0.0031	0.1404	1.0000			
MES	-0.0026	0.0198	0.0241	0.0271	0.0056	0.0024	0.0005	0.0006	0.0056	-0.0002	1.0000		
ΔCoVaR	-0.0020	0.1665	0.1509	0.1582	0.0188	0.0081	0.0086	0.0102	-0.0137	-0.0299	0.0056	1.0000	
SRISK	0.0003	0.0452	0.0488	0.0456	0.0020	0.0016	0.0002	0.0010	0.0019	-0.0003	-0.0002	0.0132	1.0000

Note: The table above shows the correlation of the regressors and the regressand (default). It can be seen that the strongest positive correlation is between equity and loans (0.9331), respectively between NPL and ROA as for the negative correlation (-0.2445). The findings make sense with regards to the high balance sheet leverage ratio, which relates the liabilities-side (i.e. equity) of the balance sheet to the assets-side (i.e. loans). High values thus suggest that there is more equity per loan than in comparison to banks with low leverage ratios. As a result, the high correlation between equity and loan mirrors the funding structure of such banks. Likewise, the return on assets is low, when the percentage of non-performing loans is high. As no payments are received by the delinquent loans, no returns are made, thus lowering the ratio relative to the assets. While the correlation is undoubtedly high, it poses no threat to the analysis, as it merely captures the structural characteristic of high equity funding in the investigated sample.

Chapter 3

Can CoCo-bonds mitigate Systemic Risk?

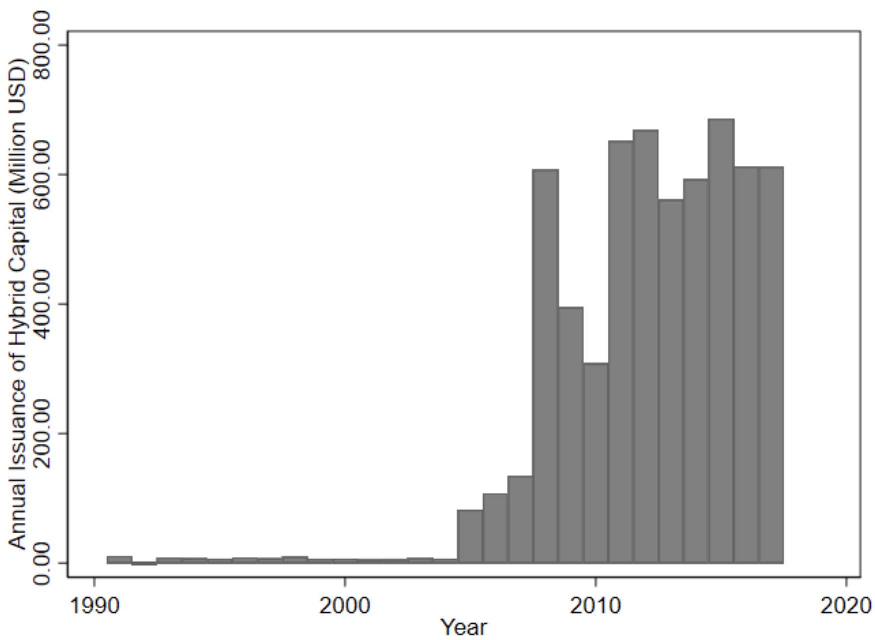
3.1 Introduction

Contingent convertible bonds (CoCo-bonds) gained particular recognition of bank regulators in the wake of the 2008 financial crisis. It exposed the vulnerability of banking systems, and the need to increase their resilience by higher quality and quantity of capital (Demirgüç-Kunt et al. (2013)). CoCo-bonds as hybrid capital instruments are predestined to serve as one contribution to this end, by combining the respective advantages of debt and equity. They are characterized as de jure debt obligations with a contractual or statutory feature to quasi-automatically convert into equity under certain conditions. While other hybrid instruments, which were predominantly used before the crisis, failed to convert into real equity instruments, the statutory conversion of CoCo-bonds allows them to provide capital when needed the most. As such, CoCo-bonds increase the resilience of the weakest link in the financial system, and hence make it more stable in its entirety.

The importance of studying hybrid capital becomes evident, when considering their growing relevance, as illustrated in Figure (3.1). It is obvious to the eye that hybrid capital has seen a steep rise in dissemination across the financial sector since the advent of the

2008 Subprime Crisis. Only the transition from Basel II to Basel III in 2010 temporarily slowed the growth in hybrid capital due to regulatory uncertainty surrounding the eligible capital tier. It has since continued its unprecedented growth at an annualized rate of almost 20 %. The new Basel Accord (i.e. Basel III) and the European Capital Requirements Regulation (CRR), respectively Capital Requirements Directive (CRD) allowed banks to cover parts of their core capital requirements by CoCo-bonds, and hence further fueled their prevalence, especially in Europe. However, despite this stellar growth, it is not undisputed, whether CoCo-bonds actually increase the resilience of banking systems. While Coffee Jr. (2011) and Avdjiev et al. (2013) find stability enhancing effects, Maes and Schoutens (2012) and Chan and Van Wijnbergen (2014) generate opposing results.

Figure 3.1: Development of annual issuance of hybrid capital over time.



We shed new light on this discussion and clarify, whether the usage of CoCo-bonds increases financial stability. Due to the plurality of proposed methods, measuring financial stability is intricate (see Gadanecz and Jayaram (2009) and Hakkio and Keeton (2009)). For the purpose of this paper, we follow the definition of Brownlees and Engle (2016) and use SRISK in order to measure a bank's impact on systemic instability. In doing so, our contribution is threefold: first, we find that CoCo-bonds do not reduce the systemic riskiness of a bank, if measured by the original SRISK formula. This finding is surprising, given the stability enhancing effect of CoCo-bonds, which constitute additional loss-absorbing capital. Our second contribution is to pinpoint this contradiction to the accounting treatment of debt CoCo-bonds in particular. Third, we propose an adjustment to the SRISK formula in order to remedy this shortcoming, and to correctly account for CoCo-bonds irrespective their treatment on the balance sheet. Using the "trigger-assumption", we imply a fictitious conversion of the CoCo-bond directly at issuance, and eliminate the undue disparities induced by differences in accounting. As a result, we can draw an unbiased picture on systemic risk, and hence financial stability. Our results are robust to different parametrizations and accounting standards, as well as issuance effects. Hence, we can make informed recommendations for policy makers and regulators alike. The necessity to do so is highlighted in unparalleled ways by the proposal of Schularick et al. (2020): they suggest to recapitalize European banks based on the SRISK measure, in order to preempt a potential capital shortfall due to heightened loan loss provisions from the Corona Crisis. However, without our adjusted SRISK formula, they will not be able to correctly identify the most vulnerable banks.

The rest of the paper is structured as follows: Section (3.2) provides the theoretical background and the relevant literature about CoCo-bonds and systemic risk. We derive our research question and hypotheses in Section (3.3). Section (3.4) summarizes our data and methodology, while Section (3.5) comprises the main results. Additional robustness tests can be found in Section (3.6), with a conclusion and an outlook given in Section (3.7).

3.2 Theoretical Background

3.2.1 CoCo-bonds

CoCo-bonds are a true subset of hybrid capital instruments. While hybrids comprise every kind of financial instrument combining features of debt and equity, not every hybrid instrument is also a CoCo-bond. Figure (3.1) illustrates the trend towards the issuance of hybrid capital instruments even before the 2008 financial crisis. Acharya et al. (2011) show that throughout the crisis a significant share of new capital issues has been in the form of hybrids, instead of common equity. Back then, Basel II allowed various different instruments to be eligible as either additional Tier 1 (AT1) or Tier 2 (T2) capital, depending on the specific national regulation. Throughout these early years, hybrids comprised preferred shares, silent participations, and various kinds of subordinated bonds broadly summarized as "innovative" hybrid capital instruments. Retrospectively, the lacking quality of some of these types of hybrids was identified as a weak-spot of the capital regulation under Basel II. Particularly, it can be argued that non-perpetual instruments or those including call options and call incentives for the issuer, interest step-up clauses, or dividend pusher clauses cannot reasonably serve as going concern Tier 1 (T1) capital. In this way, Benczur et al. (2017) note that under Basel II the true amount of banks' loss-absorbing capital was much lower than the officially reported values. Basel III raises the required quality of the financial instruments and restricts eligibility as AT1 capital to certain CoCo-bonds. In contrast to simple convertible bonds, CoCo-bonds imply for neither the issuer, nor the investor an option to convert into equity. Instead, conversion becomes mandatory if one or more contractual threshold is reached, or if the regulator considers the bank to be at the point of non-viability (PONV-trigger).

The design of CoCo-bonds varies significantly in practice with two generic types of CoCo-bonds being prevalent depending on their respective loss-absorption mechanism. In case of a breach of a pre-defined trigger threshold, the principal amount is either written down (PWD) or the financial instrument is converted into equity (C2E). More specifically, the conversion yields Common Equity Tier 1 (CET1), and hence addresses previous shortcomings under Basel II, which provided capital with questionable quality (BCBS (2010)). In this way, they are predestined to provide going concern capital to a bank under financial distress. Although important, the conversion mechanism is not exclusively decisive in determining whether the financial instrument is accounted for as debt or equity. The balance sheet treatment, however, depends critically on the accounting standards, and on the specific design of the instrument. Design features concerning the conversion price or ratio, permanent or temporary write down, or the possibility of a write up of the principal amount are left to contractual freedom. However, for regulatory eligibility as AT1, CoCo-bonds must fulfill several criteria regarding their quality to serve as going-concern capital determined by Basel III. Amongst other, the trigger must be based on the bank's regulatory CET1-capital, and amount to at least 5.125 % of the total risk-weighted assets (RWA). The exact threshold has been subject to lengthy debate. As Hart and Zingales (2011) show, some CoCo-bonds preceding the Subprime Crisis had trigger levels that were never met. While Fiordelisi et al. (2019) document a more sensible approach to the trigger levels of recently issued CoCo-bonds, they point to the instance of Banco Popular, where the CoCo-bonds still failed to convert in a timely manner. Nevertheless, CoCo-bonds are predestined to be designed in compliance with the AT1-capital requirements, because they are the only hybrid instrument that is eligible as going concern capital under Basel III. If one or more of the aforementioned criteria are not met, CoCo-bonds might still be counted towards the T2-capital. Cahn and Kenadjian (2014) provide a general overview of the regulation of CoCo-bonds according to Basel III and the European implementation through CRR and CRD IV.

The existing literature on CoCo-bonds addresses four central areas: their design, pricing, risk-taking incentives, and implications for financial stability. The conceptualization of CoCo-bonds as going concern capital goes back to the seminal work of Flannery (2005), who initially calls them "reverse convertible debenture" and later extends them to "contingent capital certificates" (see Flannery (2016)). These bonds automatically convert into common stock if a bank violates a pre-defined capital ratio, which is not based on regulatory, but book equity. In opposition to this capital ratio trigger, Raviv (2004) proposes "debt-for-equity-swaps", which are triggered if a pre-specified asset value threshold is reached. Rather than considering bank-specific trigger mechanisms, Kashyap et al. (2008) proposes a "capital insurance", ensuring that banks are recapitalized if the banking sector on aggregate reaches a situation of financial distress. In line with this idea, the Squam Lake Working Group (2009) propose a “regulatory hybrid security”, that is initially a debt-instrument, which converts into equity, if the issuing bank is under financial distress. More recently, Hart and Zingales (2011) discuss the idea of CoCo-bonds that behave like a margin account and are triggered based on CDS-spreads. A comprehensive literature review on CoCo-bonds is provided by Flannery (2014).

Although the idea of CoCo-bonds precedes the Subprime Crisis, interest in them grew manifoldly from 2008 on, in a quest for tools to strengthen the stability of the banking system. CoCo-bonds provide two channels through which bank stability can be increased. First, the coupon retention, where interest payments are deferred in order to stabilize the bank’s capital base and ease the liquidity drain. Second, the conversion, through which the de jure debt instrument becomes equity, and increases the loss-absorbing capacity. Whether, and how such a conversion affects a bank’s balance sheet equity and debt, depends on the conversion mechanism and ratio, as well as the accounting treatment. Exemplary, if a PWD-CoCo accounted for as equity is triggered, it decreases equity, but simultaneously yields the bank an extraordinary gain equal to the amount that was initially written down. At the same time, the triggering of a C2E-CoCo accounted for as debt, decreases debt and increases book equity.

Considering the effects of CoCo-bonds on the financial health of individual banks, Avdjiev et al. (2015, 2020) empirically investigate the implications of CoCo-issuances on individual bank stability. By looking at the CDS-spreads of the issuing bank, they find that banks with CoCo-bonds become more resilient. Their results thus point to an interdependence, which might be problematic for the proposal of Hart and Zingales (2011), who suggest a trigger based on the issuing bank’s CDS spread. In contrast to this bank-individual view, our study contributes to the literature on financial stability from a systemic perspective. In this way, we investigate the implications of CoCo-bonds for systemic risk and proneness to financial distress of banking systems as a whole.

Extant theoretical literature provides multiple perspectives on the relationship between the usage of CoCo-bonds and systemic risk. Avdjiev et al. (2013) postulate that the potential of CoCo-bonds to strengthen the resilience of the banking system depends in particular on their capacity to reduce systemic risk. Coffee Jr. (2011) considers contingent capital converting into equity as an effective response to systemic risk complementing regulatory supervision. Proposing a dilutive conversion of CoCo-bonds into senior shares, however, could incentivize banks to sell-off certain illiquid assets during financial crises, which would be detrimental to financial stability. Maes and Schoutens (2012) remark that CoCo-bonds could increase systemic risk, if massive investments of insurance companies in CoCo-bonds create a contagion channel from the banking to the insurance sector. Boermans and van Wijnbergen (2018) can alleviate this concern, by showing that only a marginal proportion of CoCo-bonds is cross-held by other banks and the insurance sector. In a similar way, Chan and Van Wijnbergen (2014) argue that although the conversion of CoCo-bonds strengthens the capital base of a bank, it may increase the probability of a bank run, and hence elevate systemic risk. They reason that conversion is a negative signal to the bank's depositors as well as a negative externality on other banks with correlated asset returns (particularly if banks hold each others CoCo-bonds). Koziol and Lawrenz (2012) theoretically investigate the impact of CoCo-bonds on the risk taking of owner-managers under incomplete contracts. They conclude that if owner-managers have discretion over the bank's business risk, CoCo-bonds bear adverse risk-taking incentives, increasing the idiosyncratic risk. In this way, CoCo-bonds rather fuel systemic instability, instead of mitigating it. Chan and Van Wijnbergen (2016) postulate that the widespread usage of CoCo-bonds increases systemic fragility because in particular PWD-CoCos and non-dilutive C2E-CoCos mean wealth transfers from debt holders to equity holders leading to incentives to inefficiently increase risk. Based on these ambiguous views on the effect on systemic risk, we empirically investigate this complex relationship. The following section elaborates on relevant measures for systemic risk and provides an overview of literature related to CoCo-bonds.

3.2.2 Systemic Risk

Systemic risk can be understood in many different ways, and the plurality of existing definitions highlights the still ongoing debate, about which understanding is correct. To the European Central Bank (ECB), systemic risk is "[...] the risk that financial instability becomes so widespread that it impairs the functioning of a financial system to the point where economic growth and welfare suffer materially." (ECB (2010a)). Contrarily, Schwarcz (2008) understands it as the risk that a local shock results in global repercussions because of interdependencies, respectively interconnections or external effects. The number of definitions is not bound to these two exemplary given, but illustrates the necessity of a classification of the literature. Notable attempts have been made by de Bandt and Hartmann (2000), FSB et al. (2009), and Bisias et al. (2012), respectively Benoit et al. (2017) most recently.

One approach brought forward by the ECB (2010a) is the systemic risk cube. It relates each dimension of the cube to an aspect of systemic risk. As such, it differentiates between the causes of systemic risk, its origin, and lastly manifestation. Regarding the causes, the systemic risk cube distinguishes exogenous and endogenous factors that trigger the systemic event, and hence lead to system-wide financial instability. They can either originate from a single bank (idiosyncratically) or from developments within the entire system (systemically). When they manifest, their impact can be sequential in the form of feedback loops, as described by Daníelsson et al. (2013), or simultaneous as prevalent in the literature on network effects (see Segoviano and Goodhart (2009), or Billio et al. (2012)).

Other definitions in the literature follow a less granular approach. Simply put, they differentiate between micro- and macro-level measures, which either assess the impact that systemic events have on individual banks, or the financial system as a whole. Notable contributions regarding the quantification on the bank level are the micro-level measures ΔCoVaR from Adrian and Brunnermeier (2016), respectively MES from Acharya et al. (2017), which has found its influence into SRISK by Brownlees and Engle (2016). At the other end, macro-level measures like CATFIN, as postulated by Allen et al. (2012), are noteworthy contributions to assessing the system-wide systemic risk. Irrespective of the applied definition, all systemic risk measures have individual strengths and weaknesses, depending on the dimension of systemic risk that is to be grasped. In the context of quantifying how CoCo-bonds contribute to systemic risk, these nuances make the difference in obtaining correct inference from the risk measures.

Gupta et al. (2018) use a Monte Carlo Simulation of banks' balance sheets in order to calculate ΔCoVaR in a network model where all CoCo-bonds are issued as debt. Their results indicate a strong reduction in ΔCoVaR along with less bank failures during the stress scenarios. These observations are especially true for so called "dual" trigger CoCo-bonds, where the conversion to equity, respectively the write down of the issued debt is not only dependent on a single criterion, e.g. the share price falling below a certain threshold, but the conjunction of (i) the share price, and (ii) exemplary profits falling below a certain threshold as well. A detailed discussion of this design feature can be found in the report of the Squam Lake Working Group (2009), McDonald (2013), and Allen and Tang (2016). While the findings of Gupta et al. (2018) appear desirable, they are subject to noteworthy critique. They make substantial oversimplifications, by not accounting for the different mechanics, if CoCo-bonds are issued as debt or equity. Hence, they draw a biased picture of their functioning. Furthermore, their argumentation that CoCo-bonds add additional liquidity in times of crises is flawed, as the regulator requires CoCo-bond capital to be fully paid in at issuance. Lastly, it is difficult to theorize a transmission channel between CoCo-bonds and ΔCoVaR, which consists of seven unrelated measures, such as the weekly returns of the real estate sector. Thus, the validity of employing this measure is questionable at best.

Our reservations towards ΔCoVaR in light of the aforementioned shortcomings are affirmed by the literature. Löffler and Raupach (2013) document substantial shortcomings of ΔCoVaR, which are confirmed by Kund (2018), who empirically tests the predictive power of different systemic risk measures. He finds ΔCoVaR to be the worst performing of all and generates evidence that substantiates the usage of SRISK by Brownlees and Engle (2016) for measuring systemic risk at the bank-level. We thus employ their definition of systemic risk, as an undercapitalization in the financial sector, which hence can no longer provide credit to the real economy. In order to quantify this funding gap, Brownlees and Engle have devised the systemic risk measure SRISK. Positive values indicate the presence of a funding gap, whereas negative values can be interpreted as resilience towards such adversities. The occurrence of the funding gap can be related to an extended market downturn, which is referred to as the Long Run Marginal Expected Shortfall (LRMES). It is calculated as the expected capital shortfall of the bank conditional on the occurrence of a systemic event (c), which is equal to a decline in asset prices of 10 % over the course of a month in the original paper. As such, SRISK can be understood, as an extension of the expected shortfall, as it relates the individual returns of bank i (r_i) to the returns of the market (r_m), and hence creates a systemic risk measure. In order to address structural differences between the banks, LRMES is adjusted for idiosyncratic risk through β, as well as time through $\sqrt{h}$. Formally, we can write the LRMES as:

$$LRMES_{i,t} = -\sqrt{h}\beta_i \mathbb{E}(r_{i,t+1}|r_{m,t+1} < c) \tag{3.1}$$

After obtaining the LRMES, it is incorporated in the calculation of SRISK by multiplying one minus LRMES times the adjusted equity ($E_{i,t}$) accounting for the regulatory capital fraction k. In accordance with Brownlees and Engle (2016) it was set to 8 % as approximated from the Basel Accords. Pursuant, the term is deducted from the product of book valued debt ($D_{i,t}$) and the regulatory capital fraction. We thus obtain:

$$SRISK_{i,t} = kD_{i,t} - (1-k)E_{i,t}(1 - LRMES_{i,t}) \tag{3.2}$$

This original definition though is problematic, if one is to assess the impact of hybrid capital, respectively CoCo-bonds on systemic risk. As discussed in Section (3.2.1) the accounting as debt or equity is tangent to the two balance sheet variables that are required to calculate SRISK, and hence pivotal to a correct calculation. Under the current formula, hybrid capital, such as CoCo-bonds, is not taken into account, which is why we propose an extension to Equation (3.2). We show in the following section, how our proposed "trigger-assumption" allows us to mimic the omitted loss absorbency of CoCo-bonds. As a result, we correctly grasp, how they narrow the height, respectively presence of a funding gap in the first place. From there, we derive testable hypotheses, which we describe and interpret in the subsequent sections.

3.3 Hypotheses

Throughout the existing literature on CoCo-bonds and systemic bank risk, different measures for systemic risk – as described above – are used. Fajardo and Mendes (2018) make an initial attempt to study implications for SRISK. First, they estimate SRISK for banks with and without CoCo-bonds and compare the number of defaulted banks in a stress scenario. Second, they study the market reactions of the announcement and the issuance of CoCo-bonds. Their study, though, has fundamental flaws. In particular, the authors falsely assume a generalized accounting treatment of CoCo-bonds as debt. In reality a substantial amount of CoCo-bonds is accounted for as equity, as illustrated in Tables (3.8) and (3.9) in the Appendix. Moreover, they fail to differentiate between C2E- and PWD-CoCos. This distinction is, however, vital, as both have very different effects on SRISK: while C2E-CoCos convert directly into CET1, PWD-CoCos yield an extraordinary gain, which needs to be retained in order to become regulatory capital.

The starting point of our analysis is the understanding that the original SRISK formula depends on a strict classification of all CoCo-bonds as either debt or equity and does, therefore, not properly account for hybrid capital instruments. If CoCo-bonds are not unanimously classified – as in our sample –, we expect contradicting results from their issuance. The effect of CoCo-bonds on systemic risk as measured by SRISK will crucially depend on the treatment on the balance sheet, which in turn depends on the specific design and the applicable accounting standards. If the CoCo-bond is accounted for as equity, SRISK decreases directly at emission. This effect stems from the immediate reduction of the potential funding gap due to the availability of additional equity. On the other hand, if CoCo-bonds are accounted for as debt, SRISK will increase at issuance. Even though CoCo-bonds are supposed to add additional loss-absorbing capacity, the treatment as debt increases or even invokes potential funding gaps at emission. Only upon conversion, such CoCo-bonds are properly reflected in the SRISK formula. At conversion, debt is reduced, and at the same time equity is added to the bank. The resulting net effect after conversion is the same as the effect of the usage of a CoCo-bond accounted for as equity. If a CoCo-bond is initially accounted for as equity, there is no additional effect on equity if conversion occurs at par. Figure (3.2) illustrates the different effects of CoCo-bonds on SRISK, based on their balance sheet treatment.

Figure 3.2: Expected implications of CoCo-bonds for SRISK.

	Equity	**Debt**
At emission	**SRISK** $\downarrow$	**SRISK** $\uparrow$
At conversion	**SRISK** $\rightarrow$	**SRISK** $\downarrow\downarrow$
Net effect	**SRISK** $\downarrow$	**SRISK** $\downarrow$

As a consequence of the identified differences, we cannot make generalized statements on the effects of CoCo-bonds on SRISK. The balance sheet treatment yields the counterintuitive effect that until conversion, CoCo-bonds, which are accounted for as debt, increase SRISK, despite raising the loss-absorbing capacity of a bank, just as equity CoCo-bonds do. Such a treatment contradicts the economic intuition, and implies an unjustified differentiation between otherwise comparable bonds, only because of their formal accounting treatment. In this way, SRISK discriminates against the usage of CoCo-bonds that are accounted for as debt. The correct treatment of CoCo-bonds in the SRISK formula is, however, relevant, as SRISK is manifold seen as a viable alternative to stress testing, and is frequently used by regulatory institutions to consider systemic stability (Pagano et al. (2014); Steffen (2014); Constâncio (2016)). In a worst case, the regulator wrongfully acts on a sound bank, due to misleading information about its contribution to systemic risk. A recent example would be the proposal of Schularick et al. (2020). Building on the original SRISK formula, we therefore differentiate between debt and equity, in order to aid the regulatory triage. We hence postulate the following related hypotheses:

Lemma 1 *SRISK is highly sensitive to the accounting treatment of CoCo-bonds, and thus does not correctly measure the systemic risk for issuing banks.*

Hypothesis 1 *The different treatment of CoCo-bonds accounted for as as debt, respectively equity leads to contradictory implications for financial stability.*

From a regulatory point of view, the treatment on the balance sheet does not have any consequences for the eligibility as regulatory AT1 or T2 capital. Therefore, from an economic and risk perspective, CoCo-bonds should not be treated differently. In particular, if we assume two otherwise identical bonds have the same capital quality, a CoCo-bond accounted for as debt should not increase SRISK, while a bond accounted for as equity reduces SRISK. Accordingly, we make the following adjustments to the original SRISK formula in order to account for the issuance of CoCo-bonds properly. First, we develop the hypothetical "trigger-assumption" that the issued CoCo-bonds are converted instantly at issuance. In this way, CoCo-bonds provide equity, irrespective of their accounting treatment prior to conversion. Alternatively, for PWD-CoCos, the principle amount is written down. Doing so adds equity in the form of extraordinary earnings and reduces the outstanding debt. Either way, CoCo-bonds are equally treated as loss-absorbing equity, irrespective of their balance sheet treatment. Second, we adjust the original SRISK formula as shown in Equation (3.3) to account for the insensitivity of CoCo-capital to LRMES. CoCo-bonds offer additional loss-absorbing capital in times of financial distress. Due to the trigger design, the capital is only provided in times of crisis and not ex ante. Consequently, the distributed capital is not depleted by the LRMES factor, which is why we have added it as a dedicated summand. Only once the CoCo-bonds have been converted into non-hybrid capital, the resulting equity becomes sensitive to LRMES. Taken together, we suggest for our adjusted SRISK formula:

$$\begin{aligned} SRISK_{i,t} =& k\Big(D_{i,t} - DebtCoCos_{i,t}\mathbb{1}(Triggered)\Big) \\ &-(1-k)\Big(\big(E_{i,t} - EquityCoCos_{i,t}\mathbb{1}(Triggered)\big)(1 - LRMES_{i,t}) \\ &+DebtCoCos_{i,t}\mathbb{1}(Triggered) + EquityCoCos_{i,t}\mathbb{1}(Triggered)\Big) \end{aligned} \tag{3.3}$$

Hypothesis 2 *If CoCo-bonds are properly incorporated in the SRISK formula, the usage of CoCo-bonds decreases SRISK, irrespective of their balance sheet treatment.*

Our study contributes to the literature on CoCo-bonds and systemic risk by investigating how the issuance of CoCo-bonds affects systemic risk. In particular, we show that the original SRISK formula fails to capture the specifics of CoCo-bonds in the context of systemic risk. As a result, we propose an adjustment to the SRISK formula to account for the differences in accounting treatment, remedying the inherent bias of the original SRISK formula. Doing so allows us to analyze the true impact of CoCo-bonds on systemic risk, irrespective of potential biases from the balance sheet treatment.

3.4 Data Set and Methodology

Our initial data set consists of 1,514 CoCo-issuances from 2010 until 2019 and depicts the entire universe as reported by Thomson Reuters Eikon. We narrow our sample down, by restricting it to the years after 2011, because CoCo-issuance prior to that is scarce, and might be biased due to the transition from Basel II to Basel III as shown in Figure (3.1). In spite of 110 issuances in 2019, we had to drop this year, due to missing accounting information, which are required in the calculation of SRISK. After adjusting for missing values, we obtain a sample of 533 CoCo-bonds, which were emitted by 126 banks from 33 countries around the world. Table (3.8) in the Appendix shows that almost three quarters of the CoCo-bonds in our sample are subject to the IFRS accounting regime (74 %). Amongst them, there appears to be a preference for AT1 CoCo-bonds, whereas the opposite is true for non-IFRS observations. This characteristic is in line with related literature. Avdjiev et al. (2020) report 55 % of the CoCo-bonds in their sample as AT1 capital, whereas the percentage is 52 % for ours. While the accounting as debt or equity is rather balanced for the CoCo-bonds from the IFRS domain, there is a strong preference for debt in non-IFRS banks. It is important to note that the T2 CoCo-bonds are classified exclusively as debt on the balance sheet. Table (3.9) shows that IFRS banks had no clear preference for C2E or PWD CoCo-bonds, whereas the prevalence of PWD is significantly higher for banks from non-IFRS countries. None of the CoCo-issuances has been called or triggered over the analyzed period. Thus, we have a continuous data set, free of a potential survivor bias from converted CoCo-bonds.

Our sample contains 45,864 bank-week observations from 2012 to 2018. We use weekly LRMES in order to account for sufficient volatility in the stock and market returns. Doing so prevents the estimated SRISK measure from being stale. However, for the regression analysis, we only incorporate the values reported in the first calendar week for two reasons. First, only then, the accounting information used for the calculation of SRISK can change. Second, due to the stationarity, the regression results would be biased by large numbers of almost identical values. As a result, our final sample consists of 882 bank-year observations.

We empirically test our hypotheses by employing a panel regression model with bank and time fixed-effects as depicted by α, respectively μ in Equation (3.4). Our regressands are specifications of SRISK with the variables of interest being the nominal amounts of debt-CoCos $\left(\text{CoCo}^{\text{Debt}}\right)$ and equity-CoCos $\left(\text{CoCo}^{\text{Equity}}\right)$. We subsequently control for well established bank specific and macro economic factors. On the bank level, we incorporate bank size using the logarithm of total assets. The capital structure is represented by the balance sheet leverage ratio (LR), while profitability is measured using the return on assets (ROA). We follow Laeven and Levine (2007) in measuring the income diversification using their ROID, which relates interest to non-interest income. On the country level we control for non-inflated GDP $\left(\mathrm{GDP}^{\mathrm{USD}}\right)$, annual GDP-growth $\left(\mathrm{GDP}^{\mathrm{Growth}}\right)$, annual consumer price inflation (CPI), and exuberant credit growth as measured by the credit to GDP ratio (C2GDP). We denote the coefficients for bank controls with β and the macro controls with γ to ease legibility. Subscript i refers to the individual bank, while c refers to the respective country. Time is indicated by t.

$$\begin{aligned} SRISK_{i,t+1} = & \beta_1 CoCo_{i,t}^{Debt} + \beta_2 CoCo_{i,t}^{Equity} + \beta_3 ln(Assets)_{i,t} + \beta_4 LR_{i,t} \\ & + \beta_5 ROA_{i,t} + \beta_6 ROID_{i,t} + \gamma_1 GDP_{c,t}^{USD} + \gamma_2 GDP_{c,t}^{Growth} \\ & + \gamma_3 CPI_{c,t} + \gamma_4 C2GDP_{c,t} + \alpha_i + \mu_t + \epsilon_{i,t} \end{aligned} \tag{3.4}$$

An overview over the variables and their sources can be found in Table (3.11) in the Appendix. Summary statistics and correlation metrics are provided in Tables (3.12) and (3.13) respectively.

We use the Wald test to generate evidence against autocorrelation. Likewise, heteroskedasticity can be rejected based on the results of the modified Wald test. Furthermore, we apply two treatments in order to address potential endogeneity. First, we address simultaneity and reverse causality concerns by using lagged values for the regressors in our analysis. Doing so reduces our sample to 756 observations from the initial 882, as 126 observations are used as lagged variables for the model calibration. A second source of endogeneity in our model might stem from the managerial leeway in structuring the CoCo-bond, such that it is either accounted for as equity or debt. This interdependence might be the case, if for example, highly leveraged or profitable banks systematically favor equity over debt CoCo-bonds. Hence, we apply the probit model from Equation (3.5) to verify the independence between the accounting of CoCo-bonds on the balance sheet and bank characteristics. The binary dependent variable y of the model assumes the value of one, when the CoCo-bond is accounted for as equity, respectively zero, if it is accounted for as debt. Φ denotes the standard inverse Gaussian link function in the probit model.

$$\mathbb{P}(y_{i,t} = 1|X = x_{i,t}) = \Phi\big(\beta_1 ln(Assets)_{i,t} + \beta_2 LR_{i,t} + \beta_3 ROA_{i,t} + \beta_4 ROID_{i,t} + \epsilon_{i,t}\big) \quad (3.5)$$

We generate evidence against the theorized source of endogeneity in Table (3.1). Our results hold for different measures of profitability and hence give credit to the transmission channels we have described in Section (3.3). We thus proceed with our actual analysis in the following section.

Table 3.1: Probit model to test for balance sheet accounting.

	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)
Size	0.0493	0.0615	0.0853	0.0419	0.0620
	(0.7140)	(0.6496)	(0.6369)	(0.8006)	(0.6597)
LR	-0.0199	0.0059	-0.0076	-0.0052	-0.0071
	(0.6625)	(0.8971)	(0.8684)	(0.9107)	(0.8781)
ROID	0.0105	-0.0007	0.0303	0.0968	0.0664
	(0.9875)	(0.9992)	(0.9659)	(0.8904)	(0.9232)
ROA	-0.3604				
	(0.1373)				
ROE		-0.0206			
		(0.1945)			
EBIT			-0.0000		
			(0.8358)		
Net Income				0.0000	
				(0.8213)	
Profitability					0.0024
					(0.9985)
N	509	509	509	509	509
BIC	510.7688	511.2992	512.9126	512.9052	512.9562

Note: The table above shows the coefficients and in parentheses the p-values of probit regressions of the accounting treatment of a bond on relevant bank characteristics. The binary dependent variable assumes the value one if the bond is accounted for as equity, respectively zero in the case of debt. Because we investigate whether or not a bank has issued CoCo-bonds, instead of the number of CoCo-bond issuances, the number of observations is lower compared to following tables. The bank specific variables considered are summarized in Table (3.11) in the Appendix. Model (5) uses a dummy variable that measures profitability. It is one, when the net income is positive, and zero otherwise. Significant determinants cannot be identified from this analysis. As a consequence, endogeneity concerns regarding the balance sheet treatment of the CoCo-bonds can be dispersed. Significance is denoted at the 5 %, 1 %, and 0.1 % level.

3.5 Results

3.5.1 Hypothesis 1

Table 3.2: Original SRISK formula.

	Model (1)	Model (2)	Model (3)
$CoCo^{Debt}$	-0.0074	0.0193	-0.0057
	(0.6664)	(0.2678)	(0.8311)
$CoCo^{Equity}$	-0.4848***	-0.3970***	-0.4157***
	(0.0000)	(0.0000)	(0.0000)
Size		984.2471	579.4413
		(0.2427)	(0.5938)
LR		793.4360***	780.2408***
		(0.0000)	(0.0000)
ROA		-161.4964	-111.7667
		(0.6378)	(0.7635)
ROID		2,777.3289	3,060.4407
		(0.1282)	(0.1251)
GDP^{USD}			-0.1011
			(0.3290)
GDP^{Growth}			153.9534
			(0.1436)
Inflation			17.9572
			(0.7339)
C2GDP			17.0252
			(0.2092)
Constant	6,603.8238***	-16,636.2333	-11,553.7127
	(0.0000)	(0.0815)	(0.2709)
N	756	696	637
R^2_w	0.1259	0.2471	0.2548

Note: The table above shows the coefficients and p-values (in parentheses) of regressions with bank and time fixed effects. The dependent variable is SRISK measuring systemic risk, calculated by the original formula. The variables of interest are $CoCo^{Debt}$ and $CoCo^{Equity}$, indicating the nominal amounts of CoCo-bonds accounted for as debt, respectively as equity. We find that only CoCo-bonds issued as equity reduce systemic risk under the formula proposed by Brownlees and Engle. All independent variables are one year lagged in order to disperse simultaneity concerns. Significance is denoted at the 5 %, 1 %, and 0.1 % level.

Table (3.2) depicts the test results of our first hypothesis that the original SRISK formula does not correctly account for the usage of CoCo-bonds. The dependent variable is SRISK as computed by the original SRISK formula. The variables of interest are the nominal amounts of debt-CoCos and equity-CoCos. Model (1) provides statistical evidence that the effect of CoCo-bond issuances is highly sensitive to the accounting treatment. While CoCo-bonds accounted for as equity reduce SRISK at issuance with high statistical significance, the issuance of CoCo-bonds accounted for as debt is notably insignificant. Given the idea of the SRISK formula, it is surprising, when two otherwise comparable CoCo-bonds provide both additional loss-absorbing capacity and regulatory capital, but generate contradicting results based exclusively on the accounting treatment. As a result, the regulator might wrongfully act on a sound bank, due to inconsistent results from the original SRISK formula. At the same time, the results confirm that the additional loss absorbency provided by CoCo-bonds accounted for as equity does indeed reduce SRISK. This result is intuitive but not trivial because indirect effects between the issuance of CoCo-bonds and the LRMES factor cannot be ruled out ex ante. Exemplary, it might be the case that the market perceives banks with CoCo-bonds to be substantially more resilient, such that the LRMES factor is less severe during an economic downturn. Also, the absent negative significance of the debt-CoCos underlines that there is more to the effect on SRISK than just the change in leverage. Therefore, our results confirm the theorized transmission channel between hybrid capital such as CoCo-bonds and systemic risk. Consequently, a closer investigation of the uncovered linkage is warranted.

Model (2) adds bank-specific covariates, which generate evidence against an omitted variable bias, as the previously significant intercept α becomes insignificant. At the same time, explanatory power is shifted towards the LR. It strongly contributes to explaining the riskiness of a bank from a systemic perspective. This observation is unsurprising, given that SRISK in essence measures the funding gap, which occurs, if the equity cannot support the total debt and liabilities, which are used synonymously in the work of Brownlees and Engle (2016). Given that both capital types constitute the LR, our results are coherent.

Model (3) additionally considers macro-economic control variables, but fails to improve the model, which attests to Model (2) being the correct specification to describe the underlying mechanics. Both models reinstate the previous results. While equity-CoCos continue to reduce systemic risk at a statistically highly significant level, the effect of debt-CoCos remains ambiguous, and statistically insignificant.

3.5.2 Hypothesis 2

Table (3.3) illustrates the test results of our second hypothesis, where we propose an adjustment that remedies the unequal treatment of identical CoCo-bonds, which only differ in their accounting. The dependent variable is SRISK computed by the adjusted SRISK formula as described in Equation (3.3). The variables of interest are the nominal amount of debt-CoCos and equity-CoCos. Model (1) provides statistical evidence that after the adjustment, both types of CoCo-bonds decrease SRISK at a highly statistically significant 99.9 % confidence-level. Therefore, our adjustments are adequate to eliminate the perverse disparities of the original SRISK formula. Now, for two otherwise equal CoCo-bonds, whose only difference is their accounting treatment, the true economic effect is revealed. The usage of both types of CoCo-bonds reduces SRISK by providing additional loss-absorbing capacity. Previous findings from Section (3.5.1) can mostly be reinstated for Models (2) and (3). The addition of bank-specific covariates in Model (2) shifts explanatory power from the intercept to the LR. At the same time, it moderates the effect size of the respective capital types. As before, there is no complementary influence from macro-economic control variables in Model (3). The robustness of the previous models is hence corroborated. Both variables of interest remain negative and statistically highly significant. Furthermore, we observe notable gains in the explanatory power of the models. A possible reason for this observation relates to the information conveyed in Tables (3.8) and (3.9) in the Appendix: two thirds of CoCo-bonds (68.48 %) are accounted for as debt, which omits their stability enhancing effect in the previous regressions.

Table 3.3: Adjusted SRISK formula.

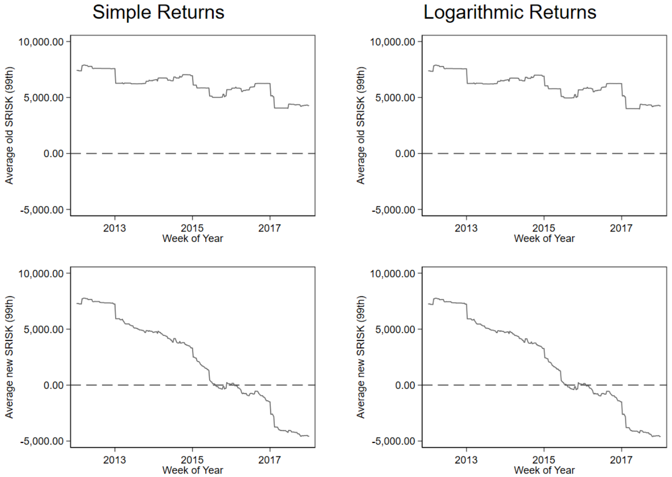
	Model (1)	Model (2)	Model (3)
$CoCo^{Debt}$	-1.0076***	-0.9806***	-1.0054***
	(0.0000)	(0.0000)	(0.0000)
$CoCo^{Equity}$	-0.4788***	-0.3906***	-0.4095***
	(0.0000)	(0.0000)	(0.0000)
Size		980.5157	601.8429
		(0.2408)	(0.5766)
LR		798.2166***	785.1944***
		(0.0000)	(0.0000)
ROA		-159.9226	-112.4378
		(0.6385)	(0.7603)
ROID		2,828.4129	3,107.2904
		(0.1184)	(0.1166)
GDP^{USD}			-0.1019
			(0.3213)
GDP^{Growth}			154.3166
			(0.1395)
Inflation			17.6509
			(0.7363)
C2GDP			16.4165
			(0.2222)
Constant	6,608.2400***	-16,689.4299	-11,769.9067
	(0.0000)	(0.0782)	(0.2583)
N	756	696	637
R^2_w	0.8518	0.8735	0.7950

Note: The table above shows the coefficients and p-values (in parentheses) of regressions with bank and time fixed effects. The dependent variable is SRISK measuring systemic risk, calculated by the adjusted formula. The variables of interest are $CoCo^{Debt}$ and $CoCo^{Equity}$, indicating the nominal amounts of CoCo-bonds accounted for as debt, respectively as equity. In contrast to Table (3.2) we find that CoCo-bonds reduce systemic risk, irrespective of their accounting. This change can be attributed to our proposed extension of the SRISK formula. All independent variables are one year lagged in order to disperse simultaneity concerns. Significance is denoted at the 5 %, 1 %, and 0.1 % level.

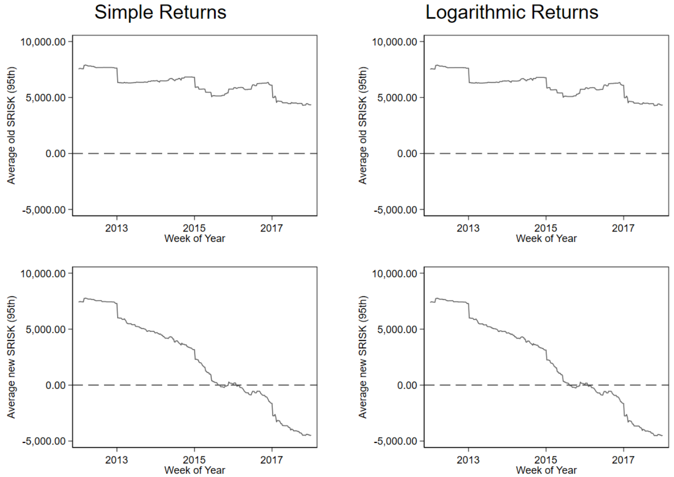
Subfigures (3.3a) and (3.3b) provide additional graphical evidence of our results, and highlight the practical implications of our findings. It can be seen in the upper row of the panel that using the original SRISK formula leads to almost unchanged levels of SRISK, in spite of CoCo-bond issuance, which de facto increases the loss-absorbing capacity of the banks and thus reduces their systemic risk ceteris paribus. It is only under our proposed adjustments in the lower row of the panel that one observes the true effect of CoCo-issuance: in line with economic theory higher levels of capitalization reduce systemic riskiness. Furthermore, we find that our adjustments indicate the absence of a funding gap from 2015 forth, as the computed average SRISK falls below zero. This observation is of paramount importance. It indicates that the regulator might wrongfully act on sound banks, because the SRISK measure in its current definition suggests a funding gap, although the opposite is true. Taken together, we show that the issuance of CoCo-bonds reduces systemic risk, if measured correctly.

Figure 3.3: Comparison between old and new SRISK measure.

(a) Subfigure 99$^{\mathrm{th}}$ Percentile



(b) Subfigure 95$^{\mathrm{th}}$ Percentile



Note: Subfigure (a) shows the difference between simple and logarithmic returns in a column-wise comparison. It is obvious to the eye, that the differences between the two return measures are marginal, and hence do not drive our results. The most interesting insight can be obtained from a row-wise comparison of the subfigure. While the top row contains the average level of SRISK under the old calculation, as depicted in Equation (3.2), the bottom row contains it with our adjustment as proposed in Equation (3.3). One directly realizes the striking difference that occurs as time progresses. Crucially, the original SRISK measure remains almost static despite the on-going issuance of additional loss-absorbing capital in the form of CoCo-bonds, and hence illustrates the problem this paper addresses. Our correction in the lower row clearly highlights that the issuance of CoCo-bonds, irrespective of their accounting treatment, reduces systemic risk. What is more, one can observe that under the new metric, SRISK on average becomes negative, which is especially interesting, given that it indicates the absence of a funding gap, whereas the top row indicates a capital shortfall. In light of this observation, the subfigure clearly illustrates the problem with the old SRISK measure, which provides a biased signal for the regulator, as it omits the loss-absorbing capacity of hybrid capital. As shown in this figure, we have remedied this shortcoming with our proposition.

Note: Subfigure (b) reinstates our findings from Subfigure (a) for a less severe market disturbance, considering the average over the worst five percent of returns, instead of the worst one percent. Again, it can be seen that our adjusted SRISK formula performs significantly better at capturing systemic risk, compared to the original formula, as we correctly capture the reduction in systemic risk that can be attributed to the issuance of additional loss-absorbing capacity in the form of CoCo-bonds. The difference between both formulas is substantial, as our adjustment generates evidence against a funding gap, illustrated by a negative SRISK from the end of 2015 forth. At the same time though, the original formula suggests that the systemic riskiness remains almost unchanged from its starting point in 2012.

3.6 Robustness

3.6.1 Parametrization

We assess the robustness of our results through a plurality of additional tests relating to the sensitivity of the parameters of the adjusted SRISK model. As such, we start by investigating the influence of different return measures on LRMES and hence SRISK. Our initial results are depicted using simple returns, and remain unchanged when using logarithmic returns, as shown in the adjacent columns of Figure (3.3). Figure (3.5) in the Appendix shows the distribution of both types of returns, and illustrates their similarities. Table (3.10) in the Appendix corroborates this characteristic by elaborating on the descriptive statistics of both return measures. While the means appear to be reasonably comparable, we have verified this numerically, applying the Wilcoxon test statistic, which indicates no differences between the two distributions.

Another driver of our results might stem from the choice of the severity of the market downturn that is used to calculate the LRMES. We have employed the most conservative estimate in our baseline results, by investigating the impact of the 99^{th} percentile of the loss distribution. Our results remain unchanged, when employing broader definitions, such as the 95^{th} percentile, as illustrated in Figures (3.3) and (3.4).

Furthermore, we winsorize the independent variables of our regression at the 1^{st} and 99^{th} percentile as a means of robustness check. Tables (3.4) and (3.5) reiterate our results, as discussed in Section (3.5), and hence disperse concerns that our results might be driven by severe outliers. While the influence of bank size becomes significant in the winsorized model, the underlying dynamics remain the same. The sign of the variables is unchanged, while their economic significance grows relative to the unrestricted models in Tables (3.2) and (3.3).

Although the results of the modified Wald test suggest homoscedasticity, we have assessed the influence of different clusters for our reported standard errors. We found no differences compared to the results in Tables (3.2) and (3.3).

Table 3.4: Original SRISK formula with winsorization.

	Model (1)	Model (2)	Model (3)
$\text{CoCo}^{\text{Debt}}$	-0.0347	-0.0172	-0.0122
	(0.0807)	(0.3964)	(0.5556)
$\text{CoCo}^{\text{Equity}}$	-0.8270***	-0.7594***	-0.7533***
	(0.0000)	(0.0000)	(0.0000)
Size		1,726.4195**	2,383.6811***
		(0.0101)	(0.0016)
LR		434.4730***	407.4773***
		(0.0000)	(0.0000)
ROA		-91.9835	-95.9905
		(0.7527)	(0.7471)
ROID		2,285.1510	2,374.9800
		(0.1340)	(0.1302)
GDP^{USD}			-0.1578*
			(0.0441)
$\text{GDP}^{\text{Growth}}$			181.2602
			(0.0759)
Inflation			6.8955
			(0.8973)
C2GDP			-2.7077
			(0.6802)
Constant	6,767.5692***	-20,148.3953**	-21,356.4799**
	(0.0000)	(0.0081)	(0.0063)
N	756	756	756
R^2_w	0.1934	0.2467	0.2541

Note: The table above shows the coefficients and p-values (in parentheses) of regressions with bank and time fixed effects on SRISK calculated by the original formula. The variables of interest are $\text{CoCo}^{\text{Debt}}$ and $\text{CoCo}^{\text{Equity}}$, indicating the nominal amounts of CoCo-bonds accounted for as debt, respectively as equity. We find that although CoCo-bonds accounted for as debt increase the loss absorbency of the issuing bank, SRISK fails to decrease. All independent variables are one year lagged in order to disperse simultaneity concerns. Our regressors are winsorized at the 1^{st} and 99^{th} percentile. Significance is denoted at the 5 %, 1 %, and 0.1 % level.

Table 3.5: Adjusted SRISK formula with winsorization.

	Model (1)	Model (2)	Model (3)
$\text{CoCo}^{\text{Debt}}$	-1.1424***	-1.1313***	-1.1295***
	(0.0000)	(0.0000)	(0.0000)
$\text{CoCo}^{\text{Equity}}$	-0.7633***	-0.7109***	-0.7038***
	(0.0000)	(0.0000)	(0.0000)
Size		1,931.0300**	2,319.1990**
		(0.0117)	(0.0070)
LR		360.9806***	343.2366***
		(0.0001)	(0.0003)
ROA		-237.0505	-214.1672
		(0.4770)	(0.5293)
ROID		1,982.3693	2,092.0286
		(0.2546)	(0.2436)
GDP^{USD}			-0.1360
			(0.1289)
$\text{GDP}^{\text{Growth}}$			171.1863
			(0.1426)
Inflation			44.3529
			(0.4680)
C2GDP			2.5344
			(0.7359)
Constant	6,808.6952***	-20,974.7265*	-21,101.8953*
	(0.0000)	(0.0138)	(0.0150)
N	756	756	756
R^2_w	0.8179	0.8261	0.8272

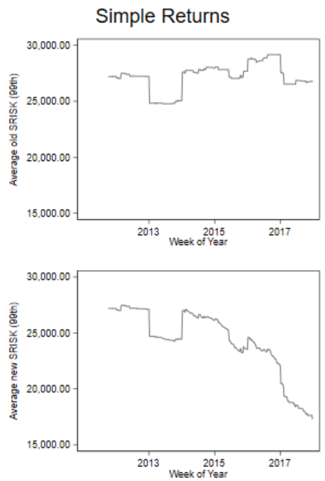
Note: The table above shows the coefficients and p-values (in parentheses) of regressions with bank and time fixed effects on SRISK calculated by the adjusted formula from Equation (3.3). The variables of interest are $\text{CoCo}^{\text{Debt}}$ and $\text{CoCo}^{\text{Equity}}$, indicating the nominal amounts of CoCo-bonds accounted for as debt, respectively as equity. We find that our adjustment remedies the inherent bias of the SRISK formula, which unduly discriminates between CoCo-bonds accounted for as debt and equity. Under our proposal, both types of CoCo-bond reduce the systemic riskiness of the issuing bank. All independent variables are one year lagged in order to disperse simultaneity concerns. Our regressors are winsorized at the 1^{st} and 99^{th} percentile. Significance is denoted at the 5 %, 1 %, and 0.1 % level.

The choice to set k to 8.00 % in the original SRISK formula, as used in Equation (3.2) and thenceforth, originates from the Pillar I requirements of Basel II. We have reapplied it to demonstrate the differences between the original SRISK formula and our methodology. In order to assess the robustness of our results, we have furthermore adjusted k to more accurately reflect the capital requirements in line with Basel III. In doing so, we accounted for two central shortcomings, compared to the work of Brownlees and Engle (2016). First, their approach uses k to relate debt to equity. However, under the cited Basel II Accord, this threshold was used to relate equity to RWA. Second, the 2008 Subprime Crisis has yielded substantial changes to the regulatory framework. Generally, the minimum equity requirements have risen from the cited 8.00 % of RWA to up to 16.50 % of RWA for global systemically important banks (G-SIBs). Moreover, the bank-specific Pillar 2 Requirement (P2R) can add additional percentage points to this ratio, and thus further supports our argument that a simple replication of the original assumptions creates a too lenient scenario. Taking these deliberations into account, we have re-evaluated Equations (3.2) and (3.3) using a k of 14.22 %.

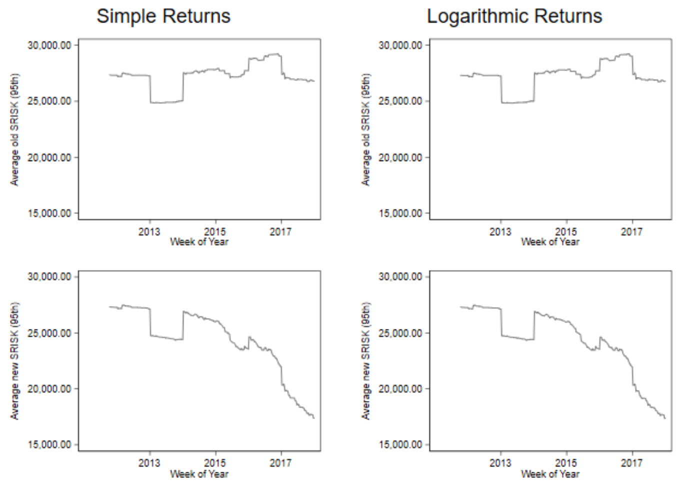
This number was obtained by dividing the median value of equity by the median value of RWA as observed in our sample. Given that the failure to adhere to the minimum capital requirements has substantial negative repercussions for the bank in question, financial institutions aim for slightly higher capital ratios, in order to safe some maneuverability. Taken together, we have created a more severe scenario, as the likelihood of a funding gap to occur has now grown, due to the larger k. The results are depicted in Figures (3.4a) and (3.4b) and show the same trend as described in Section (3.5). Our amended SRISK measure continues to decline with new issuances of CoCo-capital. At the same time, the old measure remains arguably static at a level of approximately 27 billion USD.

Figure 3.4: Comparison of alternative values of k.

(a) Comparison of SRISK with Simple and Logarithmic Returns at the 99$^{\text{th}}$ Percentile computed with an alternative k



(b) Comparison of SRISK with Simple and Logarithmic Returns at the 95$^{\text{th}}$ Percentile computed with an alternative k



Note: The figure above reinstates the robustness of our results, as has Figure (3.3a) before. We have changed the capital requirement k from 8.00 % as in the original paper to 14.22 % as we would obtain it from the data in our sample. This adjustment constitutes a more severe scenario, as a higher value of k makes the occurrence of a funding gap more likely (recall Equation (3.2)). We find that this alternation does not lead to negative SRISK values under our new formula. Nevertheless, it correctly grasps the reduction in systemic risk that can be attributed to the issuance of additional CoCo-bonds.

Note: The figure above reinstates the findings made in Subfigure (3.4a). Changing the severity of the market downturn, as we have done between Figures (3.3a) and (3.3b) with the old k, does not drive our results with the new k, as indicated by the absence of noteworthy differences.

3.6.2 Accounting Regime

By and large, the design features of a CoCo-bond are subject to contractual freedom. They can thus be chosen such that they best meet the banks' requirements. The specific design, however, determines the classification of the CoCo-bond as either debt or equity on the balance sheet. As we have shown, this attribution can have negative repercussions. If the design features necessitate a recognition of the CoCo-bond as debt on the balance sheet, the perception of systemic riskiness can be systemically biased on the bank-level. Hence, the classification as either debt or equity is a focal point in our analyses. In the interest of robustness, we demonstrated in Section (3.4) that bank characteristics do not determine whether a CoCo-bond is accounted for as debt or equity. In this section, we shed further light on the accounting standards (i.e. IFRS versus non-IFRS) as a superordinate classification criterion. Given that they are predetermined and cannot be influenced by the bank management, they might induce a bias, if comparable CoCo-bonds were systemically different recognized on the balance sheet under the respective accounting regime.

Recalling Table (3.8) from the Appendix attests to this concern, as there are statistically significant structural differences between the applied accounting standards. While non-IFRS banks issue more gone concern T2 capital, the opposite is true for IFRS banks. At the same time, there is a strong tendency for debt accounting of CoCo-bonds for non-IFRS banks, whereas the picture is less clear for IFRS banks. In light of these observations, we investigate, whether the accounting standards affect the classification as debt or equity and thus open up a transmission channel into systemic risk. Due to the invariableness of accounting regimes, we cannot use an intuitive accounting dummy in our fixed-effects regression to examine the impact of this observation (Mundlak (1978)). Instead, we resort to a decomposition of our variables of interest. Table (3.6) does not only differentiate between debt and equity CoCo-bonds, but also whether they are accounted for under (non) IFRS principles.

Table 3.6: SRISK: Comparison of old and new formula by accounting standard.

	Old Model	New Model
IFRS × $CoCo^{Debt}$	-0.0269	-1.0270***
	(0.3969)	(0.0000)
IFRS × $CoCo^{Equity}$	-0.3107***	-0.3044***
	(0.0000)	(0.0000)
(1 − IFRS) × $CoCo^{Debt}$	0.0672	-0.9318***
	(0.0951)	(0.0000)
(1 − IFRS) × $CoCo^{Equity}$	-1.9508***	-1.9465***
	(0.0000)	(0.0000)
Size	407.6885	429.0487
	(0.7020)	(0.6845)
LR	817.6635***	822.9857***
	(0.0000)	(0.0000)
ROA	-176.9095	-178.2291
	(0.6215)	(0.6157)
ROID	3,104.8911	3,151.4481
	(0.1044)	(0.0963)
GDP^{USD}	-0.0429	-0.0435
	(0.6685)	(0.6610)
GDP^{Growth}	195.1402	195.6363
	(0.0559)	(0.0531)
Inflation	22.0391	21.7142
	(0.6768)	(0.6785)
C2GDP	35.2477**	34.6560**
	(0.0080)	(0.0085)
Constant	-15,304.8313	-15,523.7485
	(0.1364)	(0.1274)
N	625	625
R^2_w	0.3367	0.8180

Note: The table above shows the coefficients and p-values (in parentheses) of regressions with bank and time fixed effects. The dependent variable is SRISK measuring systemic risk, calculated by the current formula in column one, and the adjusted formula in column two. We interact $CoCo^{Debt}$ and $CoCo^{Equity}$ with the respective accounting regimes, in order to investigate possible influences from the accounting regime. Given that we can reinstate previous results, we can curtail our results to the theorized transmission channel. Furthermore, we can disperse concerns regarding a mechanical link between the adjusted formula and our results, as the coefficients of interest are of different effect sizes. All independent variables are one year lagged in order to disperse simultaneity concerns. Significance is denoted at the 5 %, 1 %, and 0.1 % level.

If the accounting standards were an omitted force in our analyses, their influence should be most pronounced in the non-IFRS coefficients, where the majority of issued CoCo-bonds is accounted for as debt. Consistent with previous results, we find that the old model in the first column fails to recognize the loss-absorbing capacity of debt-CoCos. It is only after our proposed correction, that the undue disparity between CoCo-bonds accounted for as debt and equity is resolved. However, a strong divergence in the magnitude of the equity-CoCo coefficient becomes apparent under the non-IFRS regime, where it appears to stronger reduce bank-level systemic risk. We know from Table (3.8) in the Appendix that there are only six instances of this particular combination of CoCo-bond and accounting regime. It may thus be the case that this observation is induced by outliers. Indeed, we find the corresponding banks to be among the worst capitalized banks in the sample. They fall up to five percentage points below the average reported capital requirements, which puts two of them in the lowest decile. From this observation, another possible transmission channel opens up: could it be the case, that the issuance of CoCo-bonds increases the perceived resilience of banks and hence reduces the volatility of the issuer's shares? If this theory were true, the LRMES coefficient would be impacted, which would explain the stronger risk reduction on the systemic level. Likewise, an alternative explanation for the insignificance of the coefficient for CoCo-bonds accounted for as debt opens up, which is why we investigate this theory in the pursuant section.

3.6.3 Issuance Effects

We conduct a difference in difference analysis, in order to evaluate, whether the stock return of banks that issued CoCo-bonds has been impacted by the issuance. As the banks in our sample receive the treatment (i.e. issue CoCo-bonds) at different times, we standardize the time dimension by indexing the weeks before and after the issuance in integer increments from zero, where positive (negative) values indicate the time after (before) the treatment. Our control group has been determined through a propensity score matching, where we use assets and equity, as proxies for size, respectively capitalization. As all banks without CoCo-bonds in Thomson Reuters' Eikon were considered, we had a large population to choose from, which explains the goodness of our matches. All of them are on the support, with the differences in the estimated probabilities being no larger than 0.1321. We match every bank from the control group only once, and minimize the difference at the treatment date, in order to obtain the most similar pairs of treatment and control banks net of a possible treatment effect. Furthermore, we can verify the assumption of parallel trends, and intuitively confirm that the treatment is irreversible, as no defaults occurred, and no CoCo-bonds were called.

Table 3.7: Difference in difference test for issuance effects.

	Difference in Difference Model
Time	-0.0012
	(0.2370)
Treatment	-0.0005
	(0.6626)
Time $\times$ Treatment	0.0001
	(0.9142)
Intercept (α)	0.0025**
	(0.0029)
N	15,709
R^2	0.0002

Note: The table above depicts the results of our difference in difference analysis, where we control for market effects that might coincide with the issuance of CoCo-bonds. We find that there is neither an economically, nor statistically significant issuance effect. The independence of CoCo-issuance and the stock returns of the issuer is underscored by the significance of the constant, which hints at other explanatory powers. Significance is denoted at the 5 %, 1 %, and 0.1 % level.

Table (3.7) shows the results of the difference in difference analysis. It appears to be the case that the returns of banks have gone down through time, as the negative coefficient of Time suggests. Likewise, banks with CoCo-bonds have lower returns, as indicated by the negative coefficient of Treatment. The difference in difference estimator of the interaction term is slightly positive, but as the other coefficients, both economically, and statistically insignificant. Thus, we conclude that there are no issuance effects that stem from CoCo-bonds, which could interfere with our measurement. Our results are in line with the results of Ammann et al. (2017), Liao et al. (2017) and Avdjiev et al. (2020), who show that the issuance of CoCo-bonds affects stock prices only for a few days, and not systemically from there forth.

3.7 Conclusion

We start this paper by raising an important issue that has not received the attention of the regulator, as need be. Since the 2008 financial crisis, the issuance of hybrid capital, with CoCo-bonds being the most prominent source of it, has seen stellar growth. Given its rising importance, it is only prudent to investigate, how this capital type impacts systemic risk. Current measures of systemic risk, are mostly built around accounting measures, and fail to differentiate between capital types except for debt and equity. As such, the widespread SRISK measure is no exception to the rule. We believe, that this failure to acknowledge more granular characteristics leads to a biased view on the actual systemic risk. Indeed, our analysis shows that systemic risk is overestimated, when employing the SRISK measure, because the loss-absorbing capacity of debt-CoCos, which are the most prevalent CoCo-bonds in our sample, is omitted. As a result, regulators might look to the wrong banks in times of crisis. Under the current calculation, certain banks may exhibit a funding gap, which suggests them to be unstable, whereas the opposite is true.

We remedy this shortcoming by proposing an alternative calculation of SRISK in Equation (3.3) in order to correctly grasp the de facto systemic risk of an individual bank. By employing the "trigger-assumption", we assume that all issued CoCo-bonds are immediately converted at issuance. In this way, we eliminate the disparities in SRISK, which are solely due to a different accounting treatment. As a result, we derive a holistic framework in which both kinds of CoCo-bonds provide additional loss-absorbing capacity. This uniform treatment is particularly justified in light of the otherwise equal regulatory treatment of CoCo-bonds. We empirically find that both, equity-CoCos as well as debt-CoCos reduce a bank's contribution to systemic risk. Moreover, our adjustments show that banks, which rely on debt-CoCos, are less systemically risky than provided by the old SRISK calculation scheme, and do not necessarily have a funding gap. Consequently, we prevent the regulator from deriving wrong conclusions due to an inconsistent metric.

Future research should reinstate our findings for an even broader population of CoCo-bonds. Moreover, the generalized assumption of the SRISK formula that all liabilities will be withdrawn in times of crises might be partially unrealistic and hence should be revisited. In particular, the implicit assumption of a homogeneous reaction of deposits and other types of short-term debt is problematic. Deposit base theory motivates that even in times of financial distress a certain volume of deposits remains permanently available. The regulatory "Net Stable Funding Ratio" accounts for these differences between various types of liabilities, considering 90 to 95 % of retail deposits to be available as means of stable funding, whereas a maximum amount of 50 % of other private short-term debt is considered stable. In this way, the SRISK formula should be adjusted to account for differences in the availability of funding sources.

3.8 Appendix

Table 3.8: CoCo-bonds by accounting standard and capital tier.

	non-IFRS		IFRS	
	AT1	T2	AT1	T2
Debt	32	98	75	160
Equity	6	0	162	0
Observations	38	98	237	160
χ^2	16.19***		184.76***	

Note: The table above provides a breakdown of CoCo-bonds' accounting treatment by regulatory capital tier and applied accounting framework. The "non-IFRS" column denotes the multitude of local accounting standards. The value of Pearson's χ^2 test can be obtained from the χ^2 row. We find that the differences between the categories (i.e. AT1 and T2 issuances) are statistically significant at the 99.9 % confidence level for both accounting standards. We address this heterogeneity in dedicated analyses.

Table 3.9: CoCo-bonds by accounting standard and characteristic.

	non-IFRS		IFRS	
	C2E	PWD	C2E	PWD
Debt	1	129	123	112
Equity	3	3	76	86
Observations	4	132	199	198
χ^2	48.69***		1.13	

Note: The table above provides a breakdown of CoCo-bonds' accounting treatment by their loss absorption mechanism and applied accounting framework. χ^2 refers to the value of the test statistic according to Pearson's χ^2 test. While we find that the differences between the categories (i.e. C2E and PWD) are statistically significant at the 99.9 % confidence level for non-IFRS banks, we cannot affirm this observation for IFRS institutions.

Figure 3.5: Histograms of different return definitions.

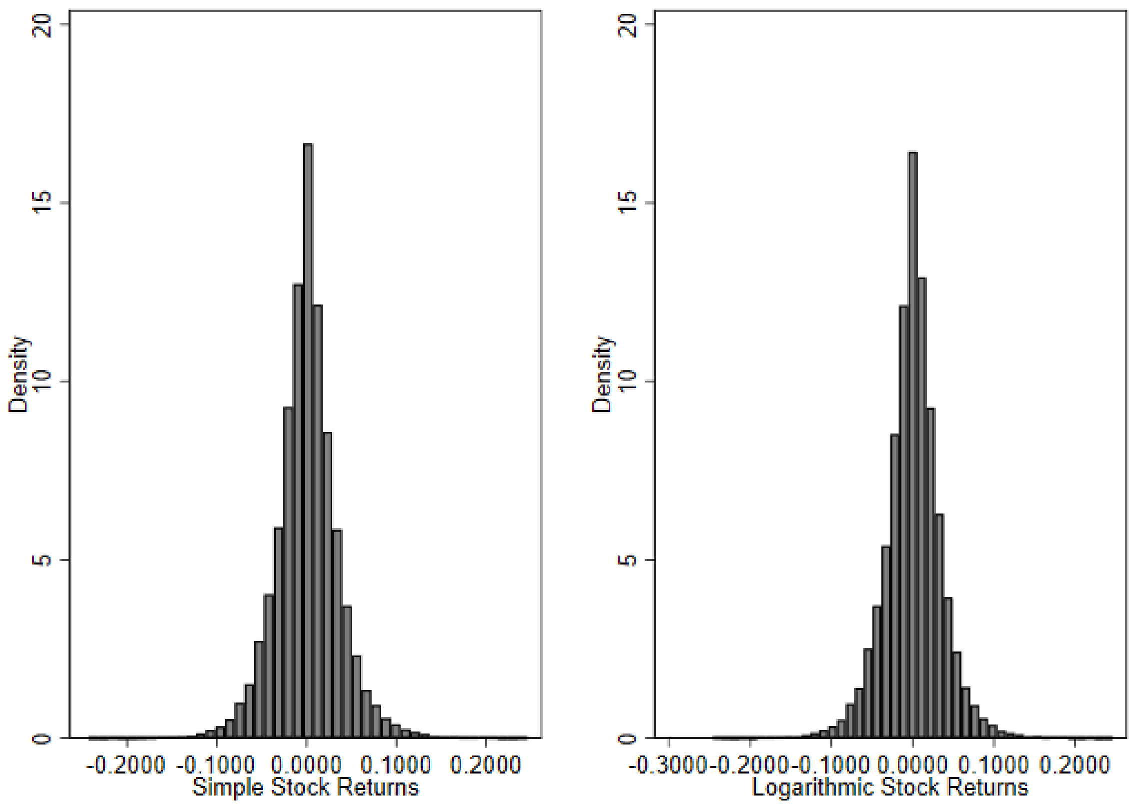


Table 3.10: Summary statistics of returns.

	N	Min	Mean	Max	Std. Dev.
simple Returns	45,862	-0.4595	0.0013	0.9298	0.0400
logarithmic Returns	45,862	-0.6152	0.0005	0.6574	0.0398

Note: As can be seen in Figure (3.5), simple returns yield slightly smaller negative values while positive values are notably larger, compared to logarithmic returns. Generally speaking, simple returns appear to be left-skewed, whereas the opposite is true for logarithmic returns. The standard deviations of both measures are comparable in terms of size. We use the Wilcoxon test to determine that there are no statistically significant differences between the two distributions.

Table 3.11: Used variables and their sources.

Variable	Description	Source
$\text{SRISK}^{\text{old}}$	SRISK as computed in Brownlees and Engle (2016)	Brownlees and Engle (2016)
$\text{SRISK}^{\text{new}}$	SRISK as computed in Equation (3.3)	Extension to the formula of Brownlees and Engle (2016)
$\text{CoCo}^{\text{Debt}}$	Nominal Amount of CoCo-bonds issued as Debt	Hand-collected from the annual report
$\text{CoCo}^{\text{Equity}}$	Nominal Amount of CoCo-bonds issued as Equity	Hand-collected from the annual report
Size	Logarithm of Total Assets	Logarithm of EIKON Item TR.TotalAssetsReported
LR	Leverage Ratio	$\frac{\text{Total Liabilities}}{\text{Total Equity}}$
ROA	Return on Assets	$\frac{\text{EBIT}}{\text{Total Assets}}$
ROID	Revenue Diversification	$1 - \left\lvert \frac{\text{Interest Income} - \text{Non Interest Income}}{\text{Interest Income} + \text{Non Interest Income}} \right\rvert$
GDP^{USD}	GDP per Capita at PPP in 2011 USD	Worldbank Indicator Code NY.GDP.PCAP.PP.KD
$\text{GDP}^{\text{Growth}}$	Annualized GDP Growth	Worldbank Indicator Code NY.GDP.MKTP.KD.ZG
Inflation	Annualized GDP Deflator	Worldbank Indicator Code NY.GDP.DEFL.KD.ZG
C2GDP	Credit to GDP	Worldbank Indicator Code FS.AST.DOMS.GD.ZS

Note: The table above outlines the data source of the used variables in this paper, and details additional calculations. We have merged multiple different data sets in order to answer our research questions. The starting point was the universe of CoCo-bonds, as reported by Thomson Reuters Eikon. From there, we amended the data set with country level macro economic control variables as reported by the Worldbank. Additional metrics have been hand-collected from the annual reports, respectively computed from the Thomson Reuters Eikon data.

Table 3.12: Descriptive statistics.

Variables	N	Min	1 %	50 %	Mean	99 %	Max	Std. Dev.
$\text{SRISK}^{\text{old}}$	40,950	-35,549.5117	-9,721.7051	400.3422	6,172.7323	66,027.8359	115,482.8047	14,611.0927
$\text{SRISK}^{\text{new}}$	40,950	-172,098.6250	-44,070.0391	78.1388	3,245.5836	63,734.8008	115,482.8047	16,084.9918
$\text{CoCo}^{\text{Debt}}$	45,864	0.0000	0.0000	0.0000	2,632.8436	56,262.7148	229,334.0156	11,566.7717
$\text{CoCo}^{\text{Equity}}$	45,864	0.0000	0.0000	0.0000	825.3359	16,530.0820	101,642.0781	4,328.9608
Size	45,864	6.5127	6.7780	11.5182	11.4008	14.6305	15.0222	1.9375
LR	45,864	3.5104	4.6905	13.2907	13.4906	27.7401	39.5339	4.9362
ROA	45,864	-2.1820	-0.0793	1.5301	1.6974	5.2637	7.4955	0.9962
ROID	42,224	0.0513	0.1272	0.6279	0.6512	1.4121	1.4950	0.3180
GDP^{USD}	45,812	4,817.1975	6,145.2946	39,700.3968	38,616.1977	90,091.4152	120,366.2801	18,857.9595
$\text{GDP}^{\text{Growth}}$	45,864	-5.7993	-2.9278	2.4492	2.9339	8.4913	25.1173	2.4998
Inflation	45,864	-25.9584	-8.8625	1.5516	1.6585	13.6501	16.5544	3.5910
C2GDP	40,872	36.0167	40.7680	165.2636	163.7235	348.6077	348.6077	61.8001

Note: This table provides summary statistics on the variables considered in the regression analysis. We display the first and ninety-ninth percentile instead of the lower and upper quartile, as we winsorize in Tables (3.4) and (3.5) in the robustness section with these percentiles.

Table 3.13: Correlation table.

Variables	$SRISK^{old}$	$SRISK^{new}$	$CoCo^{Debt}$	$CoCo^{Equity}$	Size	LR	ROA	ROID	GDP^{USD}	GDP^{Growth}	Inflation	C2GDP
$SRISK^{old}$	1.0000											
$SRISK^{new}$	0.7723	1.0000										
$CoCo^{Debt}$	0.3013	-0.3730	1.0000									
$CoCo^{Equity}$	0.1922	0.1710	0.0336	1.0000								
Size	0.5708	0.3783	0.2483	0.2491	1.0000							
LR	0.6802	0.5659	0.1164	0.0708	0.5650	1.0000						
ROA	-0.2628	-0.2025	-0.0759	-0.1115	-0.2397	-0.4277	1.0000					
ROID	0.3314	0.2667	0.0834	0.1400	0.4095	0.2830	-0.3137	1.0000				
GDP^{USD}	-0.0333	-0.0578	0.0381	0.0038	-0.0529	-0.0739	-0.2407	0.1985	1.0000			
GDP^{Growth}	-0.1491	-0.1151	-0.0413	-0.0462	-0.0615	-0.1979	0.3355	-0.3607	-0.3208	1.0000		
Inflation	-0.0755	-0.0553	-0.0318	0.0017	-0.0579	-0.1089	0.2067	-0.1223	-0.2422	0.1080	1.0000	
C2GDP	0.2751	0.2184	0.0955	0.0120	0.3514	0.4013	-0.3355	0.3050	-0.1105	-0.2686	-0.1344	1.0000

Note: This table provides pairwise Bravais-Pearson correlation coefficients of the variables included in the regression model. The highest positive coefficient can be found for the pair of the original and the new SRISK formula. This observation is unsurprising, given the similarities between the two metrics. Because no model uses both variables at the same time, this observation is unproblematic from an econometric point of view. The highest negative correlation can be attributed to the pair of $CoCo^{Debt}$ and $SRISK^{new}$. Again, this observation is in line with theory, as one expects SRISK to decrease, when CoCo capital is issued. Taken together, none of the correlations is excessive or in surprising instances, which is why we assess the probability of multicollinearity to be low.

Chapter 4

Does IFRS 9 increase Financial Stability?

4.1 Introduction

In retrospect, the 2008 Subprime Crisis revealed fundamental drawbacks in the incurred loss accounting of IAS 39 (Barth and Landsman (2010); Gebhardt (2016); Hashim et al. (2016)). Particularly criticized for its late and incomplete recognition of impairments ("too little, too late"), regulators around the globe have called for changes (G20 (2009); BCBS (2015b)). Responding to this criticism, the International Accounting Standards Board (IASB) urged a comprehensive revision of the accounting standard for financial instruments, which culminated in the release of IFRS 9 (IASB (2014b)). It constitutes a paradigm shift in the calculation of impairments for financial institutions by recognizing deteriorating credit quality in an expected credit loss (ECL) instead of an incurred loss model.

Where impairments were previously only realized when a loss event had been identified (IAS 39.59), IFRS 9 introduces a forward looking staging model, which gradually realizes them over time (IFRS 9.5.5). This adjustment is intended to lessen the severity of sudden jumps in losses ("cliff-effect"), and to diminish procyclicality. That is the positive correlation between the economic cycle and the lending activity of banks. As a result, banks have excessive capital during the expansion, while they have a shortfall during the contraction (Daníelsson (2019)). The changes from IFRS 9 are expected to address these concerns, and to increase financial stability, for which only broad definitions exist (Gadanecz and Jayaram (2009); Hakkio and Keeton (2009)). For the purpose of this paper, we look at the interaction between capital adequacy and probability of default (PD) on the bank-level, in order to quantify financial stability.

Despite its expected positive implications for financial stability, the introduction of IFRS 9 exerts influence beyond a reduction of the "cliff-effect". The earlier recognition of impairments induces a significant "front-loading" of credit losses, which likely impedes banks' ability to retain earnings. As they are a key component of Common Equity Tier 1 (CET1), not only banks' balance sheet equity, but also their regulatory capital presumably decreases. While this reduction may be desirable during the economic expansion in order to limit procyclical lending, it constitutes a noteworthy drawback for financial stability in the downturn. An impact study by the European Banking Authority (EBA) estimates an additional need for capital of 47 basis points of CET1 on average (EBA (2018d)), which translates to EUR 5.7 billion for the banks in the stress test. Another issue was raised by Abad and Suarez (2017), who analyze a portfolio of European corporate loans. They find that the impact of IFRS 9 will be most pronounced during an economic downturn, questioning the idea of reducing procyclicality as theorized by Beatty and Liao (2014).

These findings raise concerns, if the new impairment model of IFRS 9 represents an appropriate response to the experiences of the last financial crisis. In the absence of historical data, we look at the European bank stress test results, which provide a first and unique opportunity to empirically investigate this research question. Moreover, they are beneficial for our identification strategy for three reasons in particular. First, they provide two macroeconomic scenarios, which enables us to assess the severity of the methodological changes. Comparing both scenarios further allows us to infer on the theorized reduction of procyclicality. Second, the assumptions of a static balance sheet and model stock isolate the effect we want to measure. Third, they provide sufficiently granular data to address our research question in detail. In doing so, we set ourselves apart from Abad and Suarez (2017) who only analyze a portfolio of European corporate loans in a model-based setting.

Our approach to the problem necessitates the unification of two strands of literature: financial accounting at the intersection of capital adequacy and stress testing. Notable contributions are made by Novotny-Farkas (2016) and Krüger et al. (2018), who investigate the interaction between the novel impairment model and capital requirements under Basel III. Despite a manifold growth of the literature on stress testing, it is yet to address the intersection this paper identifies. Two major branches of the literature on stress testing can be discussed. One concerns stress testing as an essential part of the Basel framework (Foglia (2009)) and discusses the development of alternative risk measurement approaches (Hanson et al. (2011); Acharya et al. (2014); Schuermann (2014)) or methodological improvements (Borio et al. (2012)). The other branch empirically assesses how the publication of stress tests results influences the market value of equity or the spread of banks' credit default swaps (CDS) (Flannery et al. (2017); Ahnert et al. (2018); Sahin et al. (2020)).

Despite valuable contributions from the literature, our research question concerning the effect of IFRS 9 on financial stability remains unanswered at large. Given the implications of financial stability for the economy, it seems appropriate to fill this research gap. We construct a panel of banks from the EBA stress test exercises from 2014 to 2018 in order to address this issue. Doing so yields a sample, in which both accounting standards are present, such that we can contrast them for substantiated inference. Our analysis shows that IFRS 9 increases impairments in the short run due to the theorized "front-loading" effect. At the same time, financial stability benefits from the reduced "cliff-effect" in the long run. Drawbacks surrounding the "cliff-effect" and its contribution to procyclicality have not been fully addressed. We hence argue to increase regulatory buffers, as called for under Pillar 1 of Basel III, in order to mitigate influences that might amplify the credit cycle.

The remainder of the paper is structured as follows. Section (4.2) provides an overview of the conceptual differences between IAS 39 and IFRS 9, and disentangles their interrelation with regulatory stress testing as conducted by the EBA. In line with it we devise hypotheses concerning the effects of IFRS 9 and elaborate on the intended tests in Section (4.3). We present the analyzed data set in Section (4.4) and show the results in Section (4.5). Section (4.6) verifies our results by means of robustness tests. This paper concludes in Section (4.7), where it also gives an outlook on future research.

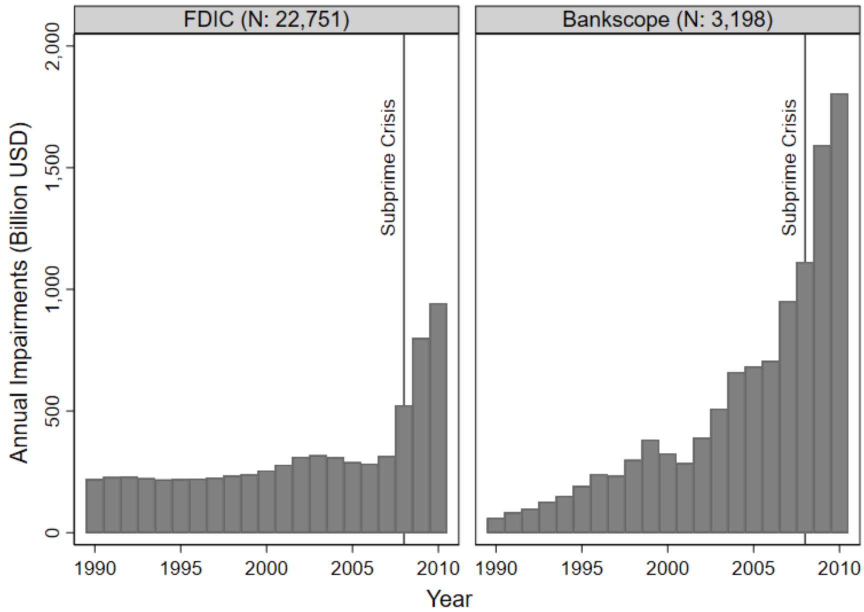
4.2 Theoretical Background

4.2.1 Differences between IAS 39 and IFRS 9

Under IAS 39, the recognition of expected losses was explicitly precluded by the standard setter. Instead, impairment losses were only incurred as of the balance sheet date, if there was objective evidence for them resulting from an event that succeeded the initial recognition of the asset (a "loss event") (IAS 39.58 f.). This definition has left plenty of leeway for judgmental factors, concerning what constitutes objective evidence (Dugan (2009)). Furthermore, it delayed the recognition of so called "day-1-losses", which occurred immediately after origination, yet were only realized as of the balance sheet date (IAS 39.AG92, IAS 39.E.4.2).

The 2008 financial crisis drew attention to this undue timely discrepancy between the loss event and its recognition (Barth and Landsman (2010); Gebhardt (2016); Hashim et al. (2016)). Moreover, the backward-looking nature of the impairment model was criticized for potentially aggravating the crisis situation (Vyas (2011); Kothari and Lester (2012); Marton and Runesson (2017)). Amongst others, the G20 raised concerns that loan loss provisioning of credit losses under the incurred loss method of IAS 39 was achieving "too little, too late" (G20 (2009); Hoogervorst (2014); BCBS (2015b)). Although Bischof et al. (2019) challenge this view, by showing that banks' loss recognition was not constrained under IAS 39, there is substantial empirical evidence concerning the negative effects of an undue delay in loan loss provisioning (Beatty and Liao (2011); Bushman and Williams (2015)). Figure (4.1) below illustrates the disparity in loan loss provisioning.

Figure 4.1: Development of impairments over time for different jurisdictions.



Based on U.S. data from the Federal Deposit Insurance Corporation (FDIC), the left graph of Figure (4.1) shows that while impairments increased around the Subprime Crisis, they only partially reflected the actual losses. The annual loss provisioning in the subsequent years exceeds that of the Subprime Crisis by a factor of almost two. The right graph of Figure (4.1) draws a similar image using global bank data from Bankscope. Again, impairments related to the last financial crisis grow twofold after the actual crisis, indicating the incomplete accounting of incurred losses. Responding to this criticism, the IASB urged a comprehensive revision of the accounting standard for financial instruments, culminating in the release of IFRS 9 (BCBS (2015a)).

With the new impairment methodology of IFRS 9 the IASB introduced a forward looking expected credit loss model (IFRS 9.5.5), requiring a more timely recognition of impairments (Landini et al. (2018)). This change was supposed to counteract the weakness of delayed credit loss recognition under IAS 39 (IFRS 9.BC.IN.2). As a consequence, the scope for the recognition of credit losses was extended beyond the static requirement of an incurred loss event as a trigger (Gebhardt (2016); Novotny-Farkas (2016)). Instead, IFRS 9 is predicated on an immediate recognition of ECL directly from a financial instrument's initial recognition (IFRS 9.5.5). The IASB defines ECL as probability-weighted estimates of credit losses (i.e., the present value of cash shortfalls) (IFRS 9.5.5.17).

Estimations of ECL shall consider all relevant information, including historical data, current conditions as well as supportable forecasts of future events and macroeconomic conditions (IFRS 9.5.5.17). Thus, IFRS 9 significantly extends the information set required to determine credit losses. The scope of the IFRS 9 impairment model includes financial assets measured at amortized cost or fair value through other comprehensive income (FVOCI). Moreover, the ECL model is applied to lease receivables, trade receivables or contract assets as well as all loan commitments and financial guarantee contracts that are not measured at fair value through profit or loss (FVPL) (IFRS 9, 4.1.2, 4.1.2a, 5.5.1, 5.5.2, BC5.118).

A key element of the IFRS 9 impairment model is the so-called three stages approach, which categorizes financial instruments according to their credit quality (i.e. 'Stage 1', 'Stage 2' and 'Stage 3'). It lessens the severity of the "cliff-effect" by gradually recognizing the ECL over the lifetime of the loan and thus reduces procyclical effects. The assignment to the stages depends on the change in credit risk since initial recognition (IASB (2013, 2014c)), and prescribes which methodology must be applied for calculating the ECL.

Stage 1 includes financial assets that were not subject to a significant increase in credit risk since initial recognition or exhibit a low credit risk as of the reporting date (IFRS 9.5.5.5). Their loss allowance is recognized as the 12-month ECL, which is defined as the share of the lifetime expected credit losses resulting from default events, which are possible within 12 months of the reporting date (IFRS 9 Appendix A). Interest revenue is calculated based on the gross carrying amount of the asset that is without deduction of the loss allowance (IFRS 9.B5.5.43).

Stage 2 includes under-performing financial assets, which exhibit a significant increase in credit risk since initial recognition. In this stage, the lifetime ECL has to be recognized (IASB (2014a); IFRS 9.5.5.3-4). It is defined as the expected credit loss from all possible default events over the expected residual life of the financial instrument (IFRS 9 Appendix A). The calculation of interest revenue remains the same as for Stage 1 (IASB (2014c); IFRS 9.5.5.3-4). At each reporting date, the reporting entities are required to evaluate whether a potentially significant increase in credit risk has occurred (IFRS 9.5.5.9). Besides the "rebuttable presumption that the credit risk on a financial asset has increased significantly since initial recognition when contractual payments are more than 30 days past due" (IFRS 9.5.5.11), the IASB provides a list of information that may be used for the assessment of a significant credit risk deterioration (IFRS 9.B5.5.17). In addition to that, the standard setter grants a "low credit risk exemption", which excludes financial assets from the continuous credit-risk assessment and allows them to remain in Stage 1, as long as they exhibit a low credit risk (IFRS 9.5.5.10). An investment grade rating by a major rating agency may serve as such an indicator (IFRS 9.B5.5.22 ff.; IFRS 9.BC5.188 f.).

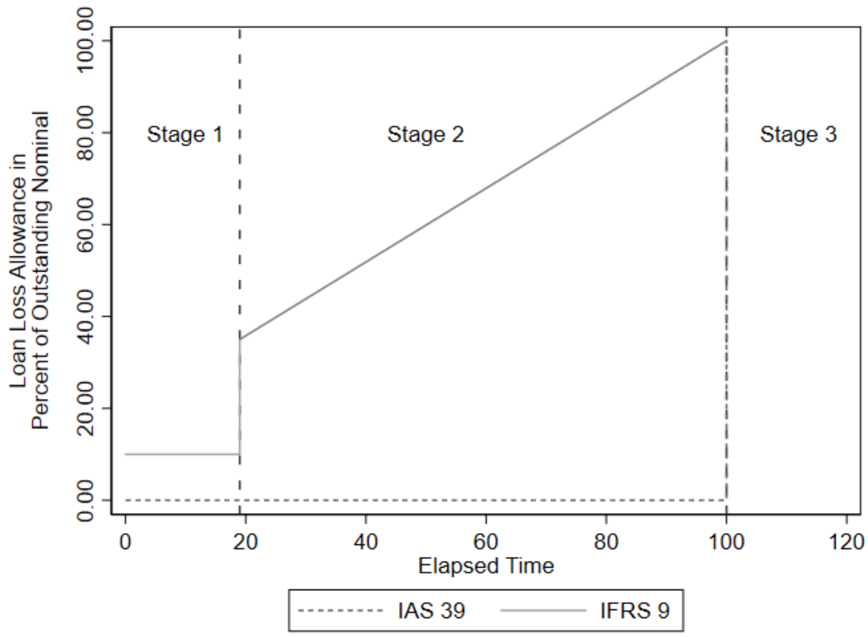
In case of a further increase in credit risk up to the status of non-performing or credit-impaired assets, the respective financial instrument must be allocated to Stage 3 (IASB (2014a)). The criteria for a financial asset to be considered as such are listed in Appendix A of IFRS 9, and largely match the objective evidences of a loss event according to the former IAS 39.59. As in Stage 2, the ECL of Stage 3 is recognized on a lifetime basis. Interest revenue is calculated based on the net carrying amount of the asset, which is the gross carrying amount less loan loss allowance (IFRS 9.5.4.1). ECL recognized in Stage 3 will likely be larger compared to Stage 2, reflecting the default position of the underlying assets. Table (4.1) provides a short overview over key implications of the three stages. A more detailed description can be found in Hartmann-Wendels et al. (2019).

Table 4.1: Stages according to IFRS 9.

	Stage 1	**Stage 2**	**Stage 3**
Classification	performing	under-performing	non-performing
Expected Loss	12 months	lifetime	lifetime
Interest Rate Calculation	gross book value	gross book value	net book value

This new impairment model appears to be a major concern for the banking industry as the initial set-up costs, as well as the adjustments to loan loss allowances are expected to increase compared to the former IAS 39 model. Since they are recognized through the P&L of the bank (IFRS 9.5.5.8), its ability to retain earnings is initially impeded (Deloitte (2013); Reitgruber et al. (2015); EBA (2016)). This interrelation negatively influences regulatory capital levels in banks (Hashim et al. (2015); Gebhardt (2016); Novotny-Farkas (2016)). Empirical evidence suggests that banks may counteract this pressure by asset sales or scaling back their loan supply with the intent to strengthen capital levels (Abad and Suarez (2017); ESRB (2017); Sánchez Serrano (2018)). However, doing so during a crisis would be diametrical to fostering financial stability, as asset prices would be further depressed and thus exacerbate the economic downturn. While the ECL model mitigates procyclicality by reducing the volatility of impairments (i.e. "cliff-effect"), it does not fully resolve the issue. Figure (4.2) illustrates the implications of a transfer from Stage 1 to Stage 2 and 3 for IFRS 9 in the full line. One can see that the reclassification from Stage 1 to Stage 2 still constitutes an abrupt increase in loan loss allowances, by transitioning from the 12 month to the lifetime ECL (Hashim et al. (2016); EBA (2016); Novotny-Farkas (2016)). While the jump in impairments is less pronounced than under IAS 39 in the dashed line, the "cliff-effect" is yet to be fully resolved. If banks offset these additional impairments by selling assets at an depressed price, it may necessitate further asset sales, starting a downward spiral as described by Brunnermeier and Pedersen (2008). The presence of countercyclical capital buffers (CCyB) as required under Pillar 1 of Basel III serves as a potential backstop against this cascade (EBA (2017); ESRB (2017)).

Figure 4.2: Illustration of the "cliff-effect" in conjunction with "front-loading".



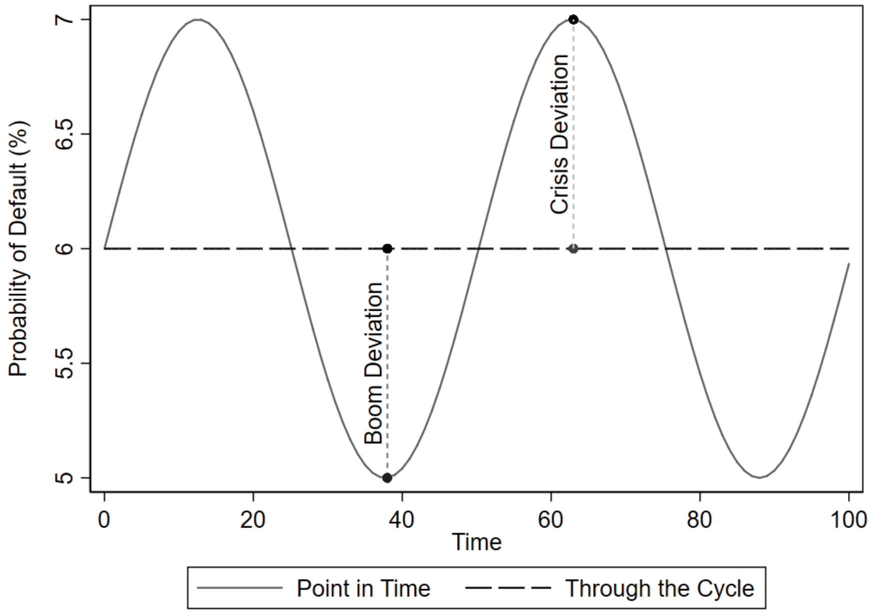
While the discussed "too little, too late" (G20 (2009); Hoogervorst (2014); BCBS (2015b)) problematic of IAS 39 has been addressed by the ECL model, not all issues of IAS 39 have been resolved (Lloyd (2018)). Another shortcoming concerns the critique that the backward-looking approach may have amplified the Subprime Crisis (Barth and Landsman (2010); Gebhardt (2016); Hashim et al. (2016)). IFRS 9 might still be susceptible to this critique because it relies on point in time (PIT) estimates for the PD. As such, only the last available data point is considered, in order to reflect the economic characteristics of the financial instrument at the reporting date (IFRS 9 BC 5.282). This approach entails profound consequences, as this individual point may be inflated during crises, respectively deflated during economic expansion (Borio and Lowe (2001)). Consequently, these estimates are subject to cyclical amplifications, and may even contribute to procyclical behavior, which is especially problematic as the PD influences the assignment to the three stages of IFRS 9 (Novotny-Farkas (2016); Vaněk et al. (2017)).

Taking this characteristic into account, the internal ratings based approach (IRB) under Basel III uses through the cycle (TTC) estimates for the calculation of the PD. The TTC approach relies on multiple historic data points, which dilutes the impact of individual points, and hence counteracts procyclicality. In light of this advantage, one wonders why the IFRS 9 ECL model does not use TTC estimates as well. The answer to this question is twofold.

First, the accounting and regulatory ECL pursue different objectives. In line with the general goal of financial reporting, the accounting ECL intends to provide useful information to the decision making of outsiders of the reporting entity (IASB (2010, 2018)). Compared to that, the regulatory ECL serves the goal of financial regulation, and as such strives to safeguard the financial system by preventing bank failures. The statistical provisioning model as described by de Lis et al. (2001) points to a viable hybrid approach, which combines the best of both worlds. Second, the usage of TTC estimates renders the stage assignment of IFRS 9 obsolete, and yields a model more similar to the U.S. current expected credit loss (CECL), where all exposures uniformly necessitate the provisioning of the expected loan losses until maturity.

Figure (4.3) illustrates the differences between the PIT and TTC approach and raises in line with our research question the concern, whether IFRS 9 has contributed to the goal of the FSF (2009) to foster financial stability by reducing the procyclical effects of IAS 39.

Figure 4.3: Illustration of the differences between TTC and PIT estimation.



Taken together, IFRS 9 presumingly reduces the "cliff-effect" by introducing a forward looking staging model. Doing so has reduced jumps in impairments, which may have procyclically enforced economic downturns. However, IFRS 9 still employs PIT instead of TTC estimators, and may thus not have gone far enough in addressing the concerns of the FSF (2009) regarding procyclicality. One way of mitigating this drawback is through designated capital buffers. Namely, the capital conservation buffer (CCB) and countercyclical capital buffer (CCyB) were designed with this intent. They amount to up to 2.5 % of the bank's risk-weighted assets (RWA). Special attention should be drawn to the CCyB, whose required paid in capital is at the discretion of national competent authorities. Out of 28 reporting countries, only one fully enforces the requirements (BIS (2018); ESRB (2019)), hence questioning their adequacy in times of crises. A more detailed discussion of the capital types and buffers can be found in Figure (4.9) in the Appendix. Another benefit of IFRS 9 concerns the more timely recognition of losses due to the ECL model. This advantage though comes at the cost of "front-loading" credit losses. Section 4.3 sheds further light on these effects and empirically assesses, whether the societal net benefit of IFRS 9 is positive.

4.2.2 Introduction to Stress Testing

Stress tests are forward-looking assessments of banks' capitalization (i.e. microprudential stress test) or the stability of the financial system as a whole (i.e. macroprudential stress test) under simulated adverse economic conditions (Hanson et al. (2011); Borio et al. (2012); Acharya et al. (2014); Ahnert et al. (2018); Duffie (2018)). One of their major objectives is to assert bank solvency (Acharya et al. (2014); Schuermann (2014)), after the last financial crisis had revealed severe (qualitative and quantitative) shortcomings in this regard (Ahnert et al. (2018)). Moreover, they facilitate supervisors to assess, whether banks comply with their regulatory capital requirements and are one tool, which European supervisors employ as part of the second pillar Supervisory Review and Evaluation Process (SREP) (BIS (2006); EBA (2018a); Paisley (2017); Ahnert et al. (2018); Riebl and Gutierrez (2018)). Additionally, regulators can test key risks such as credit, market, and liquidity risks under predefined stress scenarios to identify potential needs for capital of individual banks or to assess systemic risks, which may compromise the financial systems' stability (Ahnert et al. (2018)). Ultimately, the final disclosure of regulatory stress testing intents to improve market discipline of financial institutions and alongside increases transparency to the market (de la Lastra and Ramón (2012); Acharya et al. (2014); EBA (2018a,b)).

The first European regulatory stress test exercises were launched in 2009 and 2010 by the Committee of European Banking Supervisors (CEBS). From 2010 on, its successor, the EBA, conducted further exercises in the year 2011, and biennially from 2014 forth. Initially, the EBA's stress tests included capital hurdle rates to assess a bank's passing or failing of the test to consider further recapitalization actions in case of a failure (Riebl and Gutierrez (2018)). In the 2014 exercise, this "pass or fail threshold" was abolished. Instead, the results henceforth served as an input to the SREP (EBA (2018a,b); Riebl and Gutierrez (2018)). The effects of the stress test scenarios on banks' capital are reported in terms of the capital ratios required by Basel III (Acharya et al. (2014); EBA (2018b)). One focal item is CET1 capital, lying at the intersection of financial accounting, which this paper discusses.

Overall, the stress test coordinated by the EBA is a comprehensive exercise undertaken in close cooperation with national and EU authorities to assess the resilience of EU banks to severe market developments (de la Lastra and Ramón (2012); EBA (2018a,b); ESRB (2018)). It is conducted as a constrained bottom-up exercise, in which the participating banks apply their own internal models to project the effects of the scenarios, but are limited to the common methodology of the EBA (EBA (2018a,b)). Furthermore, it is conducted at the highest level of consolidation (i.e. group level) to assess the resilience of the largest EU banks to a (simulated) common macroeconomic baseline as well as adverse scenario over a period of three years. While there is no severely adverse scenario, as in the stress tests of the Federal Reserve, the adverse scenario of the EBA methodology can be ranked in between the adverse and severely adverse scenario of the Federal Reserve (Haselmann and Wahrenburg (2018); EBA (2018a)). Because of other divergent assumptions, such as a dynamic balance sheet, a general comparability between the two stress tests is not given. The EBA is responsible for the development of a common methodology, which all examined banks have to adhere. Furthermore, it collects the final data and disseminates it to the public to foster transparency. In devising the methodology, it is aided by the Directorate General for Economic and Financial Affairs of the European Commission, which provides the baseline scenario. The European Systemic Risk Board (ESRB) is responsible for developing the adverse macroeconomic scenarios (EBA (2018b)), while scenarios for Norwegian banks are developed by the local central bank (Norges Bank) in conjunction with the Financial Supervisory Authority of Norway (Finanstilsynet).

In November 2017 the EBA published its final methodology for the current 2018 stress test, which was launched in conjunction with the release of the macroeconomic scenarios on 31st January 2018. It lays out predefined exogenous shocks to four macroeconomic variables, such as gross domestic product (GDP) and consumer price inflation (CPI). As in previous iterations, the bottom-up exercise is subject to strict constraints. The methodological note specifies to conduct the stress test on a static balance sheet. This assumption mandates a replacement of assets and liabilities that mature during the exercises' time horizon "with similar financial instruments in terms of type, currency, credit quality at date of maturity, and original maturity as at the start of the exercise" (EBA (2018c)). In relation to the static balance sheet assumption, the EBA stress test interdicts the incorporation of anticipated capital increases by means of raises or conversions (EBA (2018a,b)). Doing so constitutes a noteworthy difference compared to other stress tests, as for example from the Bank of England, which allows capital actions (BOE (2016)). In order to gain a higher degree of transparency and comparability among banks, it is moreover assumed that participating banks maintain the same business mix and model throughout the time horizon. Ultimately, banks are subject to a model stock and can only use the internal models they have devised at the beginning of the simulation (EBA (2018c)).

For the estimation of the capital and P&L impact, the credit risk stress testing framework covers only amortized cost positions and explicitly excludes FVOCI and FVPL positions from the estimation of credit risk losses (EBA (2018a)). Especially the new impairment model of IFRS 9 implicated profound adjustments to the stress test credit risk methodology. These adjustments, which partly diverge from IFRS 9 requirements, largely concern the single scenario assumption and perfect foresight as well as the stage definitions and transfer specifications.

Under the single scenario assumption, the EBA requires banks to calculate the ECL for the baseline, respectively adverse macroeconomic scenario using only the predefined economic scenario from the regulator, instead of multiple probability-weighted cases (IFRS 9.5.5.17 (a)). Furthermore, it is assumed that banks know the precise development of the macroeconomic scenarios when calculating the lifetime ECL. It implies that all loan loss provisions for Stage 2 and Stage 3 exposures are accrued in 2018. Provisions in the following years will only be due to stage migration (EBA (2018c)). While the bidirectional transfer between Stages 1 and 2 is allowed, cures from Stage 3 are prohibited (EBA (2018a)). As under IFRS 9.5.5.5, financial instruments, whose credit risk has not increased significantly since initial recognition, are allocated to Stage 1. In line with

IFRS 9, the criterion of a significant increase in credit risk (SICR) serves as a transfer criterion to Stage 2. The methodological note clarifies that the same classification criteria may be used as under the IFRS 9 model. Furthermore, the EBA defined an additional SICR-trigger, which transfers exposures with a threefold increase over their initial lifetime PD to Stage 2. Similar to IFRS 9, a low credit risk exemption may be applied. However, the EBA specification diverges from IFRS 9 requirements, as the threshold is independent of a credit-rating. Instead, an instrument can be considered to exhibit a low credit risk, if its probability to move from Stage 1 to Stage 3 within 12 months is less than 0.3 %. Finally, exposures are allocated to Stage 3, if their credit quality decreases further to the point that it is either considered to be credit-impaired as defined under IFRS 9, defaulted as per Art. 178 of the capital requirements regulation (CRR) or classified as non-performing as per EBA Implementing Technical Standard. Banks are permitted to apply their own internal accounting practices and definitions as long as they yield more conservative results (EBA (2018a); Riebl and Gutierrez (2018)).

4.3 Hypotheses and Evaluation Methodology

The previous chapter has covered the theoretical background of the two accounting standards extensively and clearly identified their differences. The introduction of gradual loss recognition under the three stages model of IFRS 9 is expected to reduce the "cliff-effect" at the cost of introducing a "front-loading" of losses. We verify these mechanics in hypothesis one and two, before investigating the conjunction of the two effects in the third hypothesis.

Hypothesis 1 *The gradual recognition of impairments under the staging model of IFRS 9 reduces the volatility of impairments over time (i.e. the "cliff-effect").*

We test this hypothesis by comparing the variance of impairments under IAS 39 and IFRS 9. If our hypothesis is correct, we expect variance heterogeneity as the variance under IFRS 9 will be lower than under IAS 39. At the same time, the "front-loading" component should reduce the potential of banks to retain earnings, which constitute amongst other paid up instruments CET1 (Art. 28 CRR). Hence, we assume that banks cannot strengthen their regulatory capital base as measured by CET1, through retained earnings and posit:

Hypothesis 2 *The "front-loading" effect impedes banks' ability to retain earnings.*

Furthermore, we investigate how the introduction of IFRS 9 has influenced the dynamics between impairments and financial stability. We focus on the transmission channel of capital adequacy and hence the likeliness of bank failure to occur. We hypothesize that the "front-loading" effect will deplete the banks' capitalization and hence increase their PD.

Hypothesis 3 *The introduction of the IFRS 9 ECL model diminishes capital adequacy through "front-loading" losses and hence increases banks' PD.*

We test this hypothesis by computing the bank-level PD using the z-Score as in Goetz (2018). In line with the seminal work of Roy (1952), our values are normally distributed. Hence, we do not apply the standardization as suggested in Laeven and Levine (2009) or Houston et al. (2010).

$$z_{i,t} = \frac{ROA_{i,t} + CA_{i,t}}{\sigma(ROA_{i,t})} \tag{4.1}$$

The nominator of the equation above consists of the return on assets (ROA) and the capital adequacy (CA), which is measured as the ratio of equity to assets. The denominator of Equation (4.1) is the standard deviation of the ROA. The subscript t denotes time, while i refers to the bank. In essence, the z-Score can be understood as a measure for the number of standard deviations by which the ROA must fall in order to deplete the bank's equity (Boyd and Runkle (1993)).

We use the z-Score as dependent variable in our subsequent fixed-effects regression model, where we investigate the impact that impairments have on our proxy for bank PD under IAS 39, and the new IFRS 9 standard. The relationship between the likelihood of bank failure and the z-Score is inverse, such that we expect a negative coefficient on our variable of interest, impairments (IMP). We standardize impairments by total assets, in order to prevent a size bias, as large banks will naturally incur more impairments. The detailed model can be obtained from Equation (4.2).

$$\begin{aligned} z_{i,t} = \beta_1 IMP_{i,t} + \underbrace{\beta_2 LR_{i,t} + \beta_3 RISKDIV_{i,t} + \beta_4 ROID_{i,t}}_{\text{bank controls}} \\ + \underbrace{\gamma_1 HPI_{c,t} + \gamma_2 CPI_{c,t} + \gamma_3 UNEMP_{c,t} + \gamma_4 GDP_{c,t}}_{\text{macro controls}} + \alpha_i + \mu_t + \epsilon_{i,t} \end{aligned} \tag{4.2}$$

We incorporate multiple explanatory variables in our model. Our control variables for bank characteristics include the regulatory leverage ratio (LR), the risk diversification (RISKDIV), and the income diversification (ROID). Controls for bank size are obsolete for two reasons: first, the static balance sheet assumption replaces maturing assets and liabilities with comparable assets and liabilities and thus keeps total assets fixed, which would make it conceptually difficult to incorporate them in a fixed-effects model. Second, the significance assumption of the EBA makes sure that only banks with assets in excess of EUR 30 billion are part of the stress test (SSM (2013)). Hence, the interquartile range of assets is rather small and has little variation in the cross section. The LR is defined as the ratio of Tier 1 capital to total assets, while RISKDIV is a Herfindahl-Hirschman-Index, where the squared sum of the respective risk category is scaled by total RWA as shown in Equation (4.3):

$$\begin{aligned} \text{RISKDIV}_{i,t} = \Big(\frac{\text{RWA(Credit Risk)}_{i,t}}{\text{RWA(Total)}_{i,t}}\Big)^2 + \Big(\frac{\text{RWA(Market Risk)}_{i,t}}{\text{RWA(Total)}_{i,t}}\Big)^2 \\ + \Big(\frac{\text{RWA(OpRisk)}_{i,t}}{\text{RWA(Total)}_{i,t}}\Big)^2 \end{aligned} \tag{4.3}$$

In order to measure the degree of income diversification (ROID), we employ the technique of Laeven and Levine (2007) and derive an index that assumes values between zero and one. It captures the distribution between net interest income (NII) and net non-interest income (NNII), relative to their sum, the total net operation income (NOPI). The higher the value, the higher the income diversification.

$$\text{ROID}_{i,t} = 1 - \left| \frac{\text{NII}_{i,t} - \text{NNII}_{i,t}}{\text{NOPI}_{i,t}} \right| \tag{4.4}$$

Our second set of control variables includes four variables from the macroeconomic scenario, whose influence is measured by γ_n. As they are on a country-level, we introduce the subscript c to differentiate between the respective countries. We include them in order to account for the different macroeconomic scenarios, as well as structural differences between the heterogeneous countries, in which the assessed banks operate. Doing so renders the usage of country-fixed effects obsolete, as they would induce multicollinearity. Furthermore, all of them influence repayment behavior and thus the likeliness of a loan to be impaired. Especially rising unemployment (UNEMP) should severely increase the probability of delinquency, respectively default, and thus negatively influence CET1. Contrarily, a high level of GDP can be associated with a sound economic environment, in which late payments or the absence of payments occur seldom. As a result, CET1 should be high, when GDP is high. The same relationship can be attested for the House Price Index (HPI). When housing prices are high, default rates should be low, as consumers can easily refinance existing loans by borrowing against the higher value of their real estate. The influence of Consumer Price Inflation (CPI) is ambiguous. Given that wages adjust in parallel to inflation, impairment rates should decrease because the debt payments on fixed interest loans become more affordable to the consumer. To the contrary, if wage growth cannot keep up with inflation, people have less available income to allocate to debt service. We thus refrain from making an a priori assumption about the possible influence of CPI. A comprehensive list of the variables can be found in Table (4.5) in the Appendix.

Since we are interested in explaining the differences of an observed bank over time, a fixed-effects model is appealing from an econometric perspective. In particular, we apply bank and time fixed-effects, which are denoted by α, respectively μ in Equation (4.2). Applying the Hausman test deems the usage of such a model appropriate. Standard errors are clustered on the bank-level in order to account for possible heteroskedasticity. We evaluate the equation four different times, for all combinations of IAS 39 and IFRS 9 and the baseline, respectively adverse scenario. We look at the estimated coefficients in order to validate our hypothesis.

We employ the eigenvalue test of Belsley (1991) to test for multicollinearity, and disperse this concern as all condition indices are below ten. We chose this test, as it performs better for fixed-effect models, and allows to conclude on the drivers of multicollinearity, unlike e.g. the variance inflation factor (VIF). Furthermore, discarding either of the variables in our model could potentially constitute an econometrically more severe endogeneity problem due to an omitted variable. We thus proceed with the initial model, as shown in Equation (4.2). Lastly, we investigate whether the variables in our panel are stationary, using the advanced Dickey-Fuller test and generate evidence against the presence of a unit root.

The proposed methodology benefits from the stress test framework. Under the static balance sheet assumption, exposures are fixed and replaced with comparable assets at maturity. Hence, there is no inference to control for. Likewise, the prohibition of changes to the business model and capital structure exclude immeasurable effects from the model. We control for the different macroeconomic scenarios by incorporating them in our estimation model. Our methodology is thus compliant with Appendix B5.5.17 (f) of IFRS 9, which stipulates that the transition between the stages of IFRS 9 can be justified by the expectation of negative economic conditions. Moreover, the model stock assumption enables us to compare IFRS 9 models as of their inception, thus depleting the model of further biases. Consequently, we argue that, ceteris paribus, deviations in the results should be attributable to the enactment of IFRS 9.

4.4 Data Set

Our data set covers all publicly available stress test results from the EBA, respectively the European Central Bank (ECB). We merge the individual results to obtain a joint data set with 43 banks from 15 different European countries. The panel consists of empirical data from 2014 until 2018, as well as forecasts until 2020. We do not intend a counterfactual analysis, but instead try to contrast IAS 39 and IFRS 9, in order to assess the implications of the change in accounting. Although earlier stress tests are available, they were not incorporated in this paper, as they only disclose whether a regulatory hurdle rate has been exceeded or not. Our full sample represents approximately 70 % of all exposures in the Eurozone and can thus be considered representative. Two notable mergers occurred during the analyzed time. Banco Santander acquired Banco Popular Español, so that the latter was dropped from our panel. Moreover, Banco Popolare - Società Cooperativa and Banco Popolare di Milano merged. Although information for Banco Popolare are included in all three stress tests, we discontinue the time series, as Banco Popolare di Milano was not subject to previous iterations of the stress test and would thus bias the results.

Because of overlapping time frames, we have two observations for the year 2016, which is included in the 2014 and 2016 stress test. Untabulated results show that these values have high reciprocal explanatory power, when regressed on another and are significant at the 99.9 % confidence level. Thus, the selection of either year does not drive our results. We use the value from the 2014 stress test, in order to keep the time series intact for as long as possible. The data set also contains information on transitory adjustments that might arise from the new accounting standard or other regulatory influences. We decided to not incorporate them in our model for two reasons. First, only a limited number of banks makes use of them. Second, if they are being used, they are negligibly small. Because the stress test is calculated for a baseline and an adverse scenario, we have two observations in the time dimension on the bank-level. We address this issue by conducting our analyses individually for the respective scenarios. The Appendix yields the descriptive statistics for the baseline scenario in Table (4.7), whereas the results for the adverse scenario can be found in Table (4.8). Both tables have been further disaggregated, with the upper panel showing IAS 39 and the lower panel depicting IFRS 9.

4.5 Results

4.5.1 Discussion of Hypothesis 1

Figure 4.4: Visualization of impairments over the analyzed time frame.

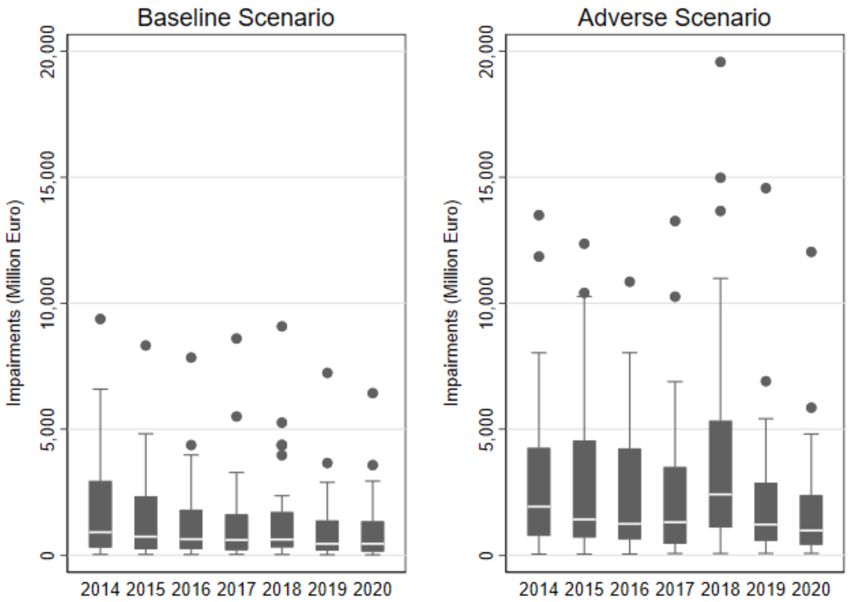


Figure (4.4) depicts the bandwidth of impairments over the analyzed stress test horizon. We chose a box-plot in order to visualize multiple dimensions of our data in an easily understandable way. The position of the 25th (75th) percentile corresponds to the lower (upper) end of the box, whereas the median is indicated by the white line within the box. The adjacent lines refer to values that are not considered outliers, as they are 1.5 times the interquartile range away from the lower and upper percentiles of our box-plot. Values exceeding this distance metric are indicated by full dots. The small box size in the baseline scenario makes it obvious to the eye, that the impairments lie narrowly together, with only little variance, as postulated by our first hypothesis. A small jump in the absolute number of outliers can be observed with the introduction of IFRS 9 at the beginning of 2018 and is in line with the theorized "front-loading" effect, which we discuss in more detail in the subsequent section. The variance under the adverse scenario is noteworthy higher. The larger body is illustrative of a wider interquartile range, which in turn further extends the adjacent lines. In accordance with our prediction, one can observe a significant reduction in volatility after the introduction of IFRS 9 in 2018, which corroborates the "front-loading" hypothesis. We proceed to empirically investigate the graphic evidence in favor of our first hypothesis by testing for variance homogeneity with Levene's test. Under our hypothesis, we expect the null hypothesis of equal variances to be rejected, as the volatilities of IAS 39 and IFRS 9 differ significantly.

Table (4.2) shows the differences between the baseline (Panel A) and adverse (Panel B) scenario for all three periods during which IFRS 9 is applicable. Using Levene's test, we calculate a test statistic in column four and computed the probability of the test statistic under variance homogeneity in column five. We find for the baseline scenario, that the initial variance homogeneity transitions into heterogeneity as time progresses. At the same time the inverse is true for the adverse scenario. We thus conclude that the impact of the new accounting standard is most pronounced under the adverse scenario, where the variances under IAS 39 and IFRS 9 converge as a result of the initial "front-loading".

Table 4.2: Comparison of variances.

	Panel A: Baseline			
	IFRS 9	ΔIAS/IFRS	Levene	Prob.
2018	1,751.49	-19.84	0.3397	0.5606
2018 – 2019	1,555.90	-215.43	3.4738	0.0635
2018 – 2020	1,449.13	-322.20	5.0408	0.0255

	Panel B: Adverse			
	IFRS 9	ΔIAS/IFRS	Levene	Prob.
2018	4,436.15	1,452.77	7.8529	0.0055
2018 – 2019	3,746.46	763.08	1.2456	0.2654
2018 – 2020	3,356.01	372.63	0.0348	0.8522

Note: The table above compares the variances of impairments under the two accounting standards. The first column depicts the length of the analyzed forecasting horizon. Columns two yields the variance of IFRS 9 over the period indicated in the first column. The third column shows the difference of the IFRS 9 values, relative to the variance we observe during the calibration period from 2014 to 2018 under IAS 39. We statistically investigate our hypothesis of variance heterogeneity by comparing Levene's test statistic and reporting the coefficient in the fourth column. Column five shows the probability of computing the value of the test statistic, under the null hypothesis of variance homogeneity. We find that the variance is statistically different in most instances. The gap widens under the baseline scenario, as indicated by the growing coefficient in column three. The negative sign suggests that the average impairments under IFRS 9 are below those of IAS 39. The opposite is true for the adverse scenario, where the initial difference is positive, but narrows down as time progresses. It suggests that banks incur more impairments under IFRS 9 than IAS 39 in the adverse scenario. These observations are in line with our hypothesis. The gradual recognition of losses under the ECL model lessens the severity of the "cliff-effect", whereas "front-loading" seems to be more dominant in the adverse scenario, and initially superimposes the decline in volatility.

4.5.2 Discussion of Hypothesis 2

Figure 4.5: Evolution of the average height of impairments.

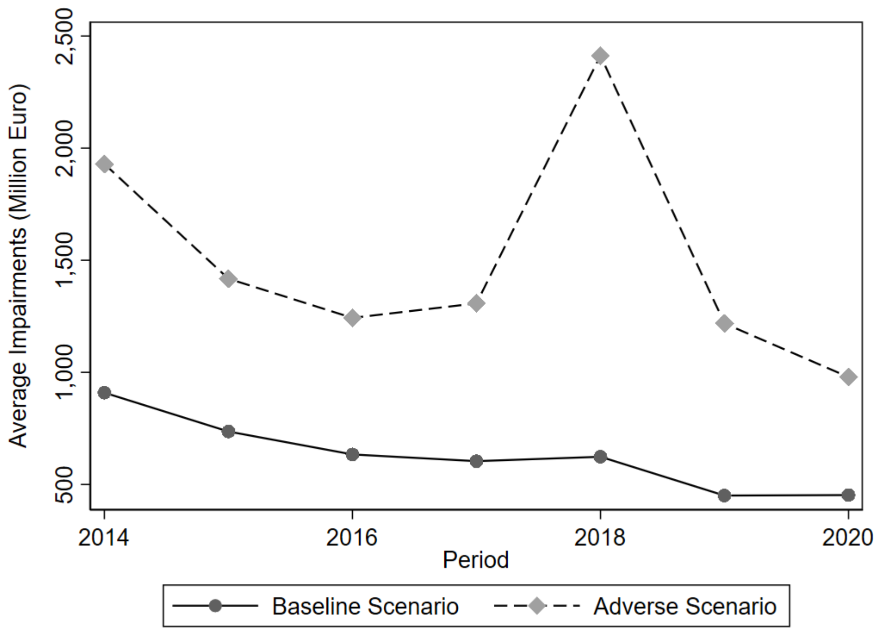


Figure (4.5) yields graphic evidence of our second hypothesis. It shows that the introduction of IFRS 9 in 2018 has coincided with a massive "front-loading" of impairments. While this observation may partially be explained by the perfect foresight approach from the stress test, it also shows that the immediate loss recognition yielded high initial impairments, yet smooths out with increasing time. In line with our second hypothesis, we proceed to empirically test the impact of this distortion on retained earnings and depict the results in Table (4.3).

Table 4.3: Comparison of average change in retained earnings.

	Panel A: Baseline			
	IAS 39	IFRS 9	Difference	Prob.
2018	875.75	1,098.74	222.99	0.3708
2018 – 2019	875.75	1,152.73	276.98	0.1452
2018 – 2020	875.75	1,164.58	288.82	0.0807

	Panel B: Adverse			
	IAS 39	IFRS 9	Difference	Prob.
2018	-517.30	-3,087.57	-2,570.27	0.0004
2018 – 2019	-517.30	-1,617.75	-1,100.48	0.0052
2018 – 2020	-517.30	-1,084.63	-567.33	0.0424

Note: The table above shows the mean change in retained earnings, under the assumption of unequal variances, in line with our insights from our first hypothesis. We compute the difference between IAS 39 and IFRS 9 by step-wise expanding the analyzed time horizon, as shown in the first column. The IAS 39 values are static, as they are the average over the period, where it was applicable (i.e. from 2014 to 2018). We find that the baseline scenario is quite optimistic, as it allows banks to increase their capital levels by retaining earnings. Surprisingly, this effect is more pronounced for IFRS 9 than IAS 39. In line with our second hypothesis, the average bank sustains losses in the adverse scenario, and hence cannot foster its capital base through retained earnings. The effect is especially strong for the first year of the analyzed horizon, which can be attributed to the discussed "front-loading". However, the longer the assessed period, the less severe the effect. This observation can be related to the gradual loss recognition, which eases the severity of initial losses over time, and is in line with our first hypothesis.

As can be inferred from the table above, the "front-loading" effect is not statistically significant for the baseline scenario. Through all analyzed time frames, banks are able to retain earnings in order to foster their capital levels. However, in case of an economic downturn, as depicted by the adverse scenario in Panel B, a very pronounced difference occurs at the onset of the crisis. Throughout the economic contraction banks are impeded in their ability to build up capital. It is only over the course of the economic contraction, that the difference narrows, and roughly vanished in the last year of observations. This finding is in line with the graphical evidence of Figure (4.5) and illustrates the severity of the “front-loading” effect, which is most pronounced during the economic downturn. We thus conclude in line with our second hypothesis that structural differences between IAS 39 and IFRS 9 exist, and that they are most pronounced at the beginning of the conversion period.

4.5.3 Discussion of Hypothesis 3

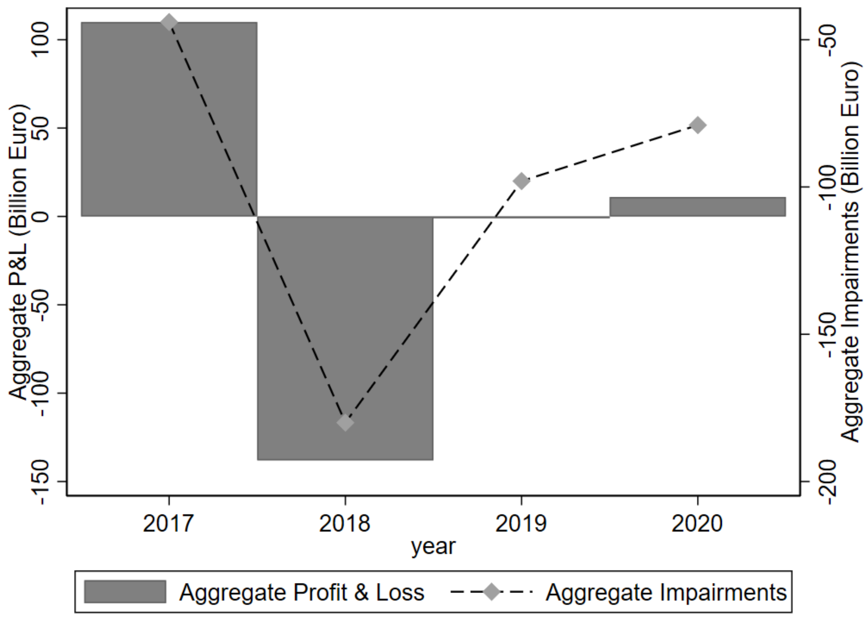
Table 4.4: Comparison of the accounting standards with y = z-Score.

	IAS 39		IFRS 9	
	Baseline	Adverse	Baseline	Adverse
IMP (%)	-1.1186**	-1.4581**	-1.5720***	-1.6033***
	(0.0048)	(0.0032)	(0.0000)	(0.0000)
LR (%)	-0.1287***	0.1545***	-0.2570	0.3279
	(0.0000)	(0.0000)	(0.0685)	(0.1682)
RISKDIV (%)	0.5936	-0.8073	-4.6357	-6.7390
	(0.7393)	(0.5381)	(0.6487)	(0.3334)
ROID ($\in \{0,1\}$)	0.0033	-0.1118	1.2683	2.4033***
	(0.9805)	(0.4236)	(0.1754)	(0.0000)
HPI (%)	0.6800	0.0206	-0.0088	0.0015
	(0.0557)	(0.0739)	(0.6412)	(0.8478)
CPI (%)	0.5273	-0.1202**	-0.0228	-0.0811
	(0.8285)	(0.0024)	(0.8023)	(0.4855)
UNEMP (%)	-0.4724	0.0013	0.0856	0.1598
	(0.9534)	(0.9580)	(0.0923)	(0.0525)
GDP (%)	0.9091	-0.1305**	-0.0647	-0.7718
	(0.8718)	(0.0025)	(0.4687)	(0.9331)
Cluster	Bank	Bank	Bank	Bank
N	172	172	129	129
R^2_{within}	0.9315	0.8912	0.4580	0.8027

Note: The table above shows the results of Equation (4.2). It can be seen that impairments (IMP) are highly significant in all models. While the importance has grown under IFRS 9, when measured in terms of the coefficient, the gap between the baseline and adverse scenario has narrowed. Taken together, the two effects yield ambiguous implications for financial stability, which we have addressed for clarification in Section (4.6). P-values are reported in parentheses. Significance is denoted at the 5 %, 1 %, and 0.1 % level.

After hypotheses one and two have confirmed the two opposing forces in terms of financial stability, our third hypothesis investigates the net impact by means of a regression analysis. We have tabulated the results in Table (4.4). They are separated by the two accounting standards, which are divided into the baseline and adverse scenario. Our findings regarding impairments are in line with our predictions. When comparing the baseline scenarios, we find that the coefficient of impairments has grown under IFRS 9. It suggests that impairments exert a stronger influence on bank PD under the new accounting standard. A possible transmission channel opens up from the theorized capital adequacy hypothesis. Due to the "front-loading" effect, banks' capitalization is negatively impacted, which in turn increases their PD as proxied through the z-Score. Figure (4.6) illustrates these deliberations by showing that banks in the adverse scenario are initially profitable in 2017, and then take a substantial hit with the introduction of IFRS 9 in the following year. This finding confirms our third hypothesis, and is in line with the results from our second hypothesis.

Figure 4.6: Aggregate impairments not measured at Fair Value through P&L.



At the same time, we find evidence in favor of the mitigation of the "cliff-effect". The gap between the baseline and adverse scenario has narrowed under IFRS 9, compared to IAS 39. As a result, banks are less vulnerable during economic downturns, as their impairments are less cyclical, and hence do no longer amplify market fluctuations. Again, this observation compliments the findings from our first hypothesis. Another notable observation concerns the leverage ratio, which is only significant under IAS 39. Our results thus suggest, that the mere importance of capitalization has been reduced under IFRS 9.

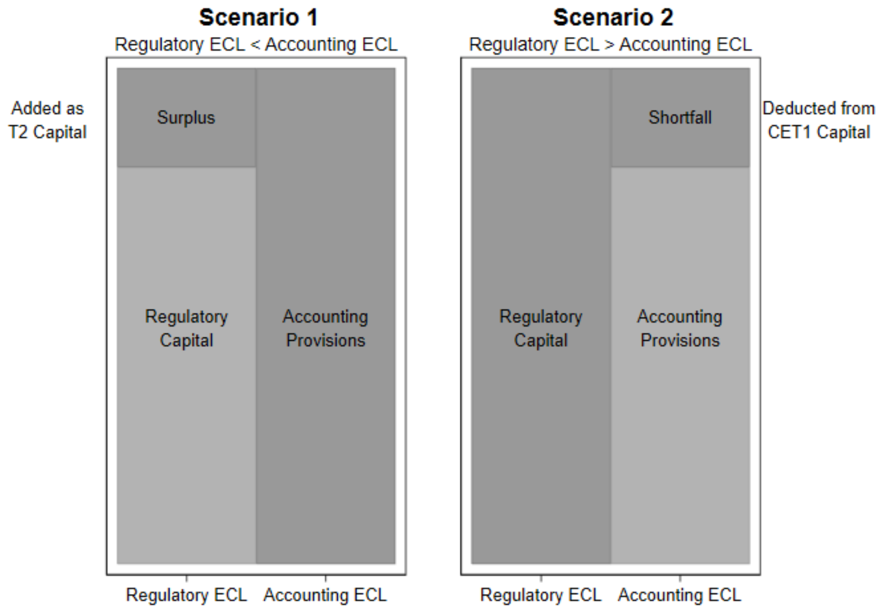
Taken together, we find that IFRS 9 has an ambiguous influence on financial stability. While undesired procyclicality in the form of the "cliff-effect" has been reduced, this advance entails the "front-loading" expected losses. Impairments thus become more important for bank stability in normal times (baseline scenario), while being less detrimental under distress (adverse scenario). Our findings complement early conjectures made by the EBA (2018b) and hint at interrelations that might influence the lending behavior of banks.

4.6 Robustness

Due to the research setting, it was not feasible to conduct some common robustness checks. We employ subsampling as part of our identification strategy in order to differentiate between the baseline and the adverse scenario. Therefore, a further disaggregation would only lead to inconclusive subsets with no meaningful data. Likewise, the limited sample size has depleted winsorization or truncation of meaning. To the contrary, the volatile observations under macroeconomic stress actually contain significant information for our research question in light of the "cliff-effect". It may seem appealing to understand the introduction of IFRS 9 as a treatment effect, and to hence employ a difference in difference approach for the identification strategy. However, since there are no banks in the stress test that are not subject to the new accounting standard, the required control group cannot be constructed. Likewise, an event study appears appealing, but is not feasible as the event is clustered around the introduction of IFRS 9 (MacKinlay (1997)).

Another approach of testing our results stems from Art. 159 of the CRR. In order to ensure consistency between regulatory and economic capital, it mandates the comparison of the calculated ECL for general and specific credit risk adjustments in line with IFRS 9 to the regulatory ECL according to the CRR. From this comparison, two scenarios can arise, as shown in Figure (4.7). Either, an ECL shortfall, when IFRS 9 provisions are short of CRR provisions, or a surplus in the reciprocal case.

Figure 4.7: Possible constellations when comparing the ECL.



Note: Additions to Tier 2 Capital can only be made up to a maximum of 0.6 % of RWA.

Under real world conditions, surpluses as in the first scenario of Figure (4.7) can be considered Tier 2 capital up to a maximum of 0.6 % of RWA. However, the methodological note of the European stress test interdicts this attribution, in order to yield more conservative results (EBA (2018c)). In line with Art. 36 (1) (d) CRR, a shortfall will be deducted from the Common Equity Tier 1 and thus relates to a section of the equity, which also contains the focal point of our analysis: retained earnings. A detailed numerical example can be found in Krüger et al. (2018), while the economic reasoning behind this approach is explained in Figure (4.10) in the Appendix. We consider our first and second hypothesis robust, if we can observe through the comparison of the regulatory and accounting ECL that IFRS 9 initially yields higher loan loss provisions than IAS 39 due to the "front-loading" effect. As a result, the number of observed shortfalls should decrease. Furthermore, we expect the nominal amount of the shortfall to lessen due to the expected loss framework.

Figure 4.8: Evolution of the ECL shortfall in the baseline and adverse scenario.

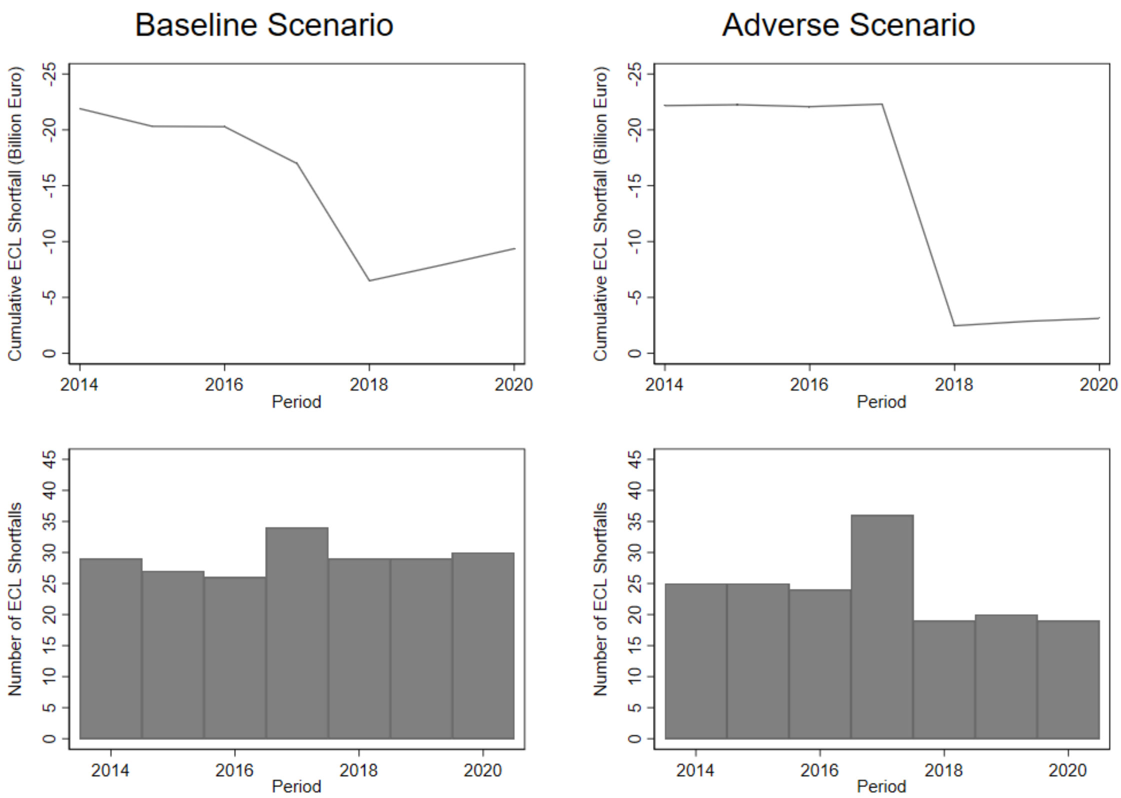


Figure (4.8) depicts the cumulative nominal shortfall in the first row of the panel, and the absolute number of shortfalls in the second row. The graphs in the first column relate to the baseline scenario, whereas the second column contains the adverse scenario. The aggregate shortfall drops sharply with the introduction of IFRS 9 in 2018. This drop can arguably be attributed to the discussed "front-loading" effect, which has increased impairments, and hence narrowed the gap between both ECL measures. We thus interpret it as further evidence for our second hypothesis. The impact is most pronounced for the adverse scenario, where a steep decline can be observed in contrast to the steady reduction under the baseline scenario. Likewise, the number of banks with an ECL shortfall is elevated for both scenarios prior to 2018, giving further credibility to the "front-loading" explanation. Furthermore, it can be seen that the number of banks with a shortfall under IFRS 9 is almost 20 % below the number reported under IAS 39. In relation to the question of reduced procyclicality, this observation might be understood as an indication that under macroeconomic stress banks are no longer subject to self-enforcing amplifications.

We furthermore challenge the robustness of our model by conducting a pseudo-treatment study, where we estimated Equation (4.2) under the assumption that the introduction of IFRS 9 did not occur in 2018, but in any other year. We find in untabulated results that the observed mechanisms are only significant for the year of the de facto introduction. Lastly, our results are valid for fully loaded, respectively transitory reported numbers.

4.7 Conclusion

This paper sets out to generate novel insights regarding the implications of the new IFRS 9 impairment model for financial stability. The shift from an incurred to an expected credit loss model has released two opposing forces, whose net effect remains ambiguous ex ante. While the more timely recognition of losses under IFRS 9 fosters financial stability by mitigating procyclical effects, it also weakens capital adequacy, potentially setting off this benefit. We investigate this impact, using the z-Score as a proxy for the likelihood of a bank to fail. It is an especially suitable measure in this context, as it emphasizes the transmission channel between capital adequacy, which is impacted by IFRS 9, and probability of default.

We posit three main hypotheses in connection with the advent of IFRS 9. First, the gradual loss recognition of the ECL model should decrease the volatility of impairments. The "cliff-effect" of the incurred loss model of former IAS 39 represented a major source of procyclicality, which should be mitigated by the gradual loss recognition under IFRS 9. Although a dampened version of the "cliff-effect" still persists in the shift from Stage 1 to Stage 2, it should be attenuated by the CCB and CCyB. Second, initial impairments under IFRS 9 should be higher compared to IAS 39 due to the earlier recognition of impairments under the ECL approach and the resulting "front-loading" effect. Third, the impact of impairments on capital adequacy, and, subsequently, on the probability of bank failure, should be the strongest at the onset of the crisis. Over the further course of the crisis, this impact should decrease.

Given the absence of historical data on IFRS 9, we draw on the ECB banking stress test results in order to investigate our hypotheses. They allow us to study the implications of the new ECL impairment model on bank resilience and financial stability based on the entire loan portfolios of major European banks in unparalleled granularity. In the absence of archival data from actual crises, the specified stress test scenarios offer a first and unique opportunity to empirically explore the implications of IFRS 9 on banks' reported results. We can investigate whether procyclicality was indeed reduced by comparing the baseline and adverse scenario of the stress test. Furthermore, all banks adhere to the same assumptions and methodologies. We could thus exclude noise from immeasurable effects and are confident to have measured the true implications of IFRS 9.

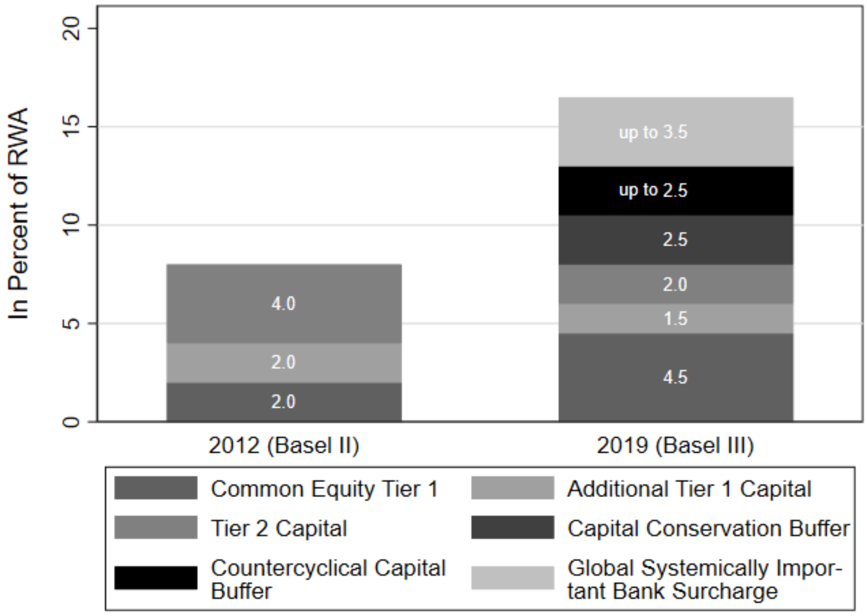
With regards to our first hypothesis, our analysis reveals that the "cliff-effect" of IAS 39 has been weakened under IFRS 9, which indicates the potential of the staging model to enhance financial stability of the banking sector in the future. We continue our investigation by assessing whether the reduction of the "cliff-effect" came at the theorized cost of "front-loading". Consistent with our second hypothesis, we find that impairments grow excessively at the beginning of the adverse scenario. However, the gap between the two accounting standards narrows as time progresses. The findings of our third hypothesis confirm the previous results. Impairments under IFRS 9 exert a stronger influence on financial stability, when proxied as banks' PD through the z-Score. The gap between an economic downturn and the status quo though has been reduced. This observation suggests that the procyclicality of impairments has been decreased, which in turn would benefit financial stability.

Although, the results of our paper indicate that the introduction of IFRS 9 has successfully diminished the severity of the "cliff-effect", this goal was achieved at the cost of "front-loading" expected credit losses. As a result, less secure loans incur higher costs at their initial recognition, which might lead to a credit supply shock from banks, and deter bank managers from acquiring such loans in the secondary market. Consequently, asset quality becomes more important under the new accounting standard. Our findings do not only concern the management of financial institutions, but can also be extended to regulatory and supervisory policy discussion. While the timelier recognition of expected credit losses under the IFRS 9 approach may have positive effects on financial stability and bank resilience, not all issues of the preceding IAS 39 have been resolved. Our results highlight the need to pay in the new regulatory capital buffers, in order to contain the remaining "cliff-effect" from Stage 1 to Stage 2 and above. Only then, the desired stabilization of the financial system will truly be achieved. The recent announcement of the German regulator to raise the CCyB to 0.25 % as of Q3 2020 can be seen as a step in this direction.

The combination of stress test results and accounting requirements opens up a plurality of new research questions. While the usage of forecasted data allows us to give an early assessment of the implications of IFRS 9, future research should try to assert our findings using actual data. Moreover, it appears prudent to repeat this study with coming stress test results, in order to increase the power of our tests. It also seems appropriate to assess how the differences between IAS 39 and IFRS 9 manifest under the standard and internal ratings based approach of the Basel Accords. Lastly, it would be advisable to compare the ECL staging model to the upcoming current expected credit loss (CECL) model of the FASB. Unlike IFRS 9, all eligible exposures are immediately recognized with their lifetime ECL under the CECL model. Doing so eliminates the "cliff-effect" and thus further reduces procyclicality, which only stems from the usage of PIT estimates under the proposal of the FASB. However, at the same time, the "front-loading" effect will be even more pronounced, necessitating a further investigation into the implications in the context of financial stability.

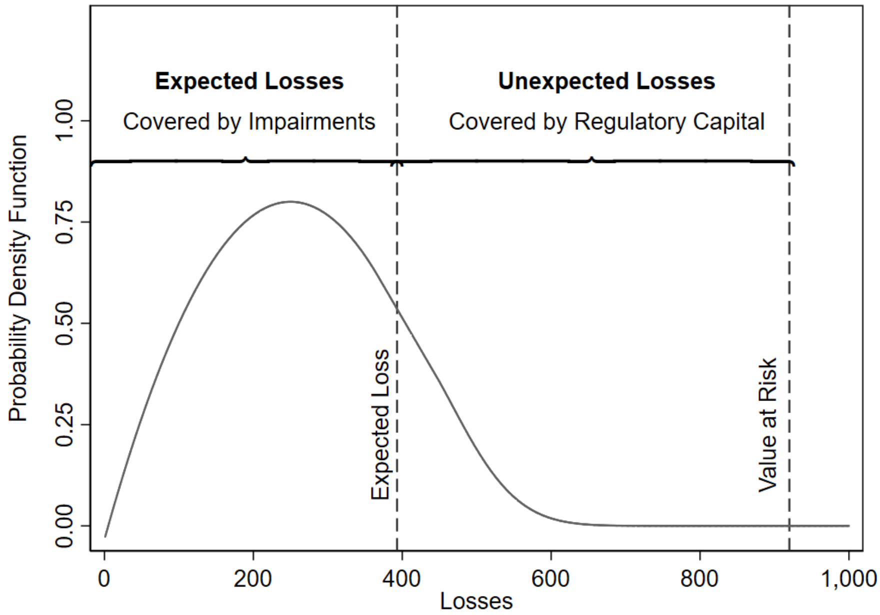
4.8 Appendix

Figure 4.9: Illustration of the differences between Basel II and Basel III.



Note: The introduction of Basel III has yielded significant changes to the own funds of banks. Not only has the composition of equity changed, but also have other items been added to the Pillar 1 requirements, in order to make banks more resilient. Large, systemically relevant banks (GSIB) for example are now subject to individual capital surcharges based on their perceived riskiness, as measured in so called buckets. A pivotal element in the context of this paper is the Countercyclical Capital Buffer (CCyB). It is intended to increase the resilience of the banking sector by means of an additional capital accumulation in periods of excessive credit growth. In downturns, when losses materialize, this buffer shall be used to mitigate impairments, reducing the risk of an extenuated credit supply constrained by regulatory capital requirements. To this day, only one out of 28 reporting countries fully enforces the requirements (BIS (2018); ESRB (2019)). The CCyB's adequacy in times of crises may consequently be questioned. Our study on the impact of the ECL model in a crisis scenario may thus be useful in the regulatory debate to actively use the additional loss-absorbing buffer and set the CCyB rate above 0.0 % to strengthen the capitalization of banks in good times.

Figure 4.10: Illustration of expected and unexpected losses.



Note: Figure (4.10) explains the economic intuition behind the ECL shortfall comparison in more detail. It was argued that two cases can occur: either a shortfall or a surplus. While the shortfall is deducted from CET1 capital, the surplus can be added as T2 capital up to a maximum of 0.6 % of RWA. The reason for this unequal treatment becomes obvious, when constructing an example. Consider the first case, a shortfall. It occurs, when the impairments do not suffice to cover the expected losses. As a result, an area between expected and unexpected losses arises, where losses are not covered by neither accounting nor regulatory capital. Hence, the deduction from CET1 to account for these losses. In the contrary case of a surplus, the unexpected losses covered by regulatory capital are also partially covered by impairments. As a result, the bank covers certain losses twofold. In order to prevent the bank from being charged twice, the idea is to offset the negative implications from this welcomed conservatism by allowing the addition of the double covered capital to T2 capital up to a maximum 0.6 % of RWA.

Table 4.5: Used variables and their sources.

Variable	Description	Source
ASSETS	Total Assets	Own Computation: ASSETS = $\frac{\text{T1 Capital}}{\text{Leverage Ratio}}$
NI	Net Income	Item 993014[1], Item 1690715[2], Item 183615[3]
ROA	Return on Assets	Own Computation: ROA = $\frac{\text{Net Income}}{\text{Total Assets}}$
z	z-Score	Own Computation: $z = \frac{\text{ROA + CA}}{\sigma(\text{ROA})}$
IMP	Amortized Impairments	Item 993007[1], Item 1690710[2], Item 183610[3]
LR	Leverage Ratio	Item 1690858[2], Item 183112[3]
RISKDIV	Risk Diversification	Own Computation: RISKDIV = $\sum_{j=1}^{3} \text{Risk } (\%)_{itj}^{2}$
NII	Net Interest Income	Item 993001[1], Item 1690701[2], Item 183601[3]
NNII	Net Non-Interest Income	Item 993002[1], Item 1690705[2], Item 183605[3]
NOPI	Net Operating Income	Item 993005[1], Item 1690709[2], Item 183609[3]
ROID	Income Diversification	Own Computation: ROID = $1 - \left\lvert \frac{\text{NII}-\text{NNII}}{\text{NOPI}} \right\rvert$
ECL	ECL Shortfall	Item 993416[1], Item 1690815[2], Item 183716[3]
HPI	Housing Price Inflation	ESRB[4]
CPI	Consumer Price Inflation	ESRB[4]
UNEMP	Unemployment Rate	ESRB[4]
GDP	Gross Domestic Product	ESRB[4]

Note: (1) as obtained from the 2014 Stress Test Results website. (2) as obtained from the 2016 Stress Test Results website. (3) as obtained from the 2018 Stress Test Results website. (4) as obtained from the macroeconomic scenario diffused by the ESRB. Total Assets for 2014 were extrapolated from the actual values, in line with the "static balance sheet" assumption of the bank stress test.

Table 4.6: Correlation table.

	z-Score	IMP (%)	LR (%)	RISKDIV ($\in \{0, 1\}$)	ROID ($\in \{0, 1\}$)	HPI (%)	CPI (%)	UNEMP (%)	GDP (%)
z-Score	1.0000								
IMP (%)	-0.1119	1.0000							
LR (%)	0.7780	0.2688	1.0000						
RISKDIV ($\in \{0, 1\}$)	0.1404	0.2272	0.2049	1.0000					
ROID ($\in \{0, 1\}$)	0.0034	-0.1754	-0.0823	0.0519	1.0000				
HPI (%)	0.3466	-0.2290	0.0754	0.0621	0.0907	1.0000			
CPI (%)	0.1571	-0.1194	0.0244	-0.0075	0.0820	0.5741	1.0000		
UNEMP (%)	-0.0923	0.2871	-0.0831	0.2429	-0.0662	-0.1984	-0.3393	1.0000	
GDP (%)	0.2235	-0.1801	0.0247	0.0594	-0.0516	0.6971	0.4627	-0.0938	1.0000

Note: The table above shows the correlations between the regressand and regressors from Equation (4.2). The dimension of the respective variable has been added in parentheses, where applicable. The strongest positive correlation can be observed between the leverage ratio and the z-Score. Given that a modified version of the leverage ratio influences the numerator of the z-Score as a measure of capital adequacy, this observation appears unproblematic. In light of the otherwise small size of the correlations, no pair raises concerns for the empirical analysis of our paper.

Table 4.7: Descriptive statistics of the baseline scenario.

Panel A: IAS 39

	Obs.	*Min.*	$Q_{0.25}$	$Q_{0.50}$	$Q_{0.75}$	*Max.*	σ
z-Score	172	-0.55	0.70	1.14	1.63	14.36	1.88
IMP	172	0.02	0.13	0.20	0.37	2.07	0.31
LR	172	1.69	4.07	4.87	6.07	24.95	15.81
RISKDIV	172	0.38	0.63	0.68	0.74	0.86	0.10
ROID	172	0.00	0.00	0.00	0.00	0.98	0.31
HPI	172	-4.30	1.50	4.00	5.60	8.70	2.98
CPI	172	0.30	1.15	1.40	1.70	2.80	0.42
UNEMP	172	3.80	5.50	7.40	10.40	25.70	4.47
GDP	172	0.20	1.50	1.85	2.40	4.50	0.69

Panel B: IFRS 9

	Obs.	*Min.*	$Q_{0.25}$	$Q_{0.50}$	$Q_{0.75}$	*Max.*	σ
z-Score	129	-0.12	0.94	1.39	1.90	3.69	0.77
IMP	129	0.01	0.09	0.13	0.23	0.99	0.18
LR	129	3.31	4.86	5.55	6.61	12.14	1.95
RISKDIV	129	0.42	0.66	0.72	0.75	0.86	0.09
ROID	129	0.02	0.47	0.65	0.91	0.99	0.27
HPI	129	-1.60	2.90	3.80	4.80	12.60	1.94
CPI	129	0.70	1.40	1.70	2.00	2.90	0.42
UNEMP	129	2.90	3.90	5.00	8.80	14.80	3.09
GDP	129	1.30	1.60	1.70	2.30	4.30	0.66

Note: The table above depicts the descriptive statistics of IAS 39 (Panel A) and IFRS 9 (Panel B) in the baseline scenario. Notable variables include the income diversification (ROID), which is highly skewed, and shows that the banking sector in the EU is highly dependent on interest income. In both panels the ROID almost assumes the maximal theoretical value of one. This observation is unsurprising, given that the examined banks are all based in bank-based economies, where companies rely on credit, to finance their operations.

Table 4.8: Descriptive statistics of the adverse scenario.

Panel A: IAS 39

	Obs.	*Min.*	$Q_{0.25}$	$Q_{0.50}$	$Q_{0.75}$	*Max.*	σ
z-Score	172	-3.49	-0.57	-0.09	0.39	14.36	2.06
IMP	172	0.03	0.31	0.40	0.83	3.53	0.57
LR	172	1.60	3.49	4.17	4.97	24.95	15.87
RISKDIV	172	0.36	0.61	0.67	0.73	0.86	0.10
ROID	172	0.00	0.00	0.00	0.76	1.00	0.31
HPI	172	-19.20	-9.90	-5.50	-3.50	9.20	4.42
CPI	172	-3.90	-0.50	0.35	0.90	2.40	1.14
UNEMP	172	4.60	7.20	9.50	11.10	26.80	4.48
GDP	172	-4.10	-1.60	-1.10	-0.70	0.90	0.79

Panel B: IFRS 9

	Obs.	*Min.*	$Q_{0.25}$	$Q_{0.50}$	$Q_{0.75}$	*Max.*	σ
z-Score	129	-3.01	-0.92	-0.16	0.36	1.96	1.08
IMP	129	0.05	0.23	0.40	0.65	2.23	0.41
LR	129	1.88	3.90	4.61	5.45	11.23	1.89
RISKDIV	129	0.40	0.65	0.70	0.75	0.86	0.09
ROID	129	0.00	0.37	0.60	0.82	1.00	0.31
HPI	129	-31.10	-11.60	-7.20	-2.40	10.00	7.92
CPI	129	-1.80	0.10	0.40	1.10	2.70	0.89
UNEMP	129	3.80	6.10	8.10	10.20	15.90	3.08
GDP	129	-31.00	-2.20	-1.20	0.00	1.90	5.48

Note: The table above shows the descriptive statistics of IAS 39 (Panel A) and IFRS 9 (Panel B) in the adverse scenario. We can reinstate the description from Table (4.7) at large. Again, the high skewness in terms of diversification characterizes the European banking market.

Bibliography

Abad, J. and Suarez, J. (2017). Assessing the Cyclical Implications of IFRS 9 – A Recursive Model. *ESRB Occasional Paper Series*.

Acharya, V., Engle, R., and Pierret, D. (2014). Testing Macroprudential Stress Tests: The Risk of Regulatory Risk Weights. *Journal of Monetary Economics*, 65:36–53.

Acharya, V. V., Gujral, I., Kulkarni, N., and Shin, H. S. (2011). Dividends and Bank Capital in the Financial Crisis of 2007-2009. Technical report, National Bureau of Economic Research.

Acharya, V. V., Pedersen, L. H., Philippon, T., and Richardson, M. (2017). Measuring Systemic Risk. *Review of Financial Studies*, 30(1):2–47.

Acharya, V. V. and Ryan, S. G. (2016). Banks' Financial Reporting and Financial System Stability. *Journal of Accounting Research*, 54(2):277–340.

Adrian, T. and Brunnermeier, M. K. (2016). CoVaR. *American Economic Review*, 106(7):1705–1741.

Ahnert, L., Vogt, P., Vonhoff, V., and Weigert, F. (2018). The Impact of Regulatory Stress Testing on Bank's Equity and CDS Performance. *Swiss Institute of Banking and Finance*.

Allen, L., Bali, T. G., and Tang, Y. (2012). Does Systemic Risk in the Financial Sector predict future Economic Downturns? *The Review of Financial Studies*, 25(10):3000–3036.

Allen, L. and Tang, Y. (2016). What's the Contingency? A Proposal for Bank Contingent Capital triggered by Systemic Risk. *Journal of Financial Stability*, 26:1–14.

Altman, E. I. (1968). Financial Rations, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *The Journal of Finance*, 23(4):589–609.

Ammann, M., Blickle, K., and Ehmann, C. (2017). Announcement Effects of Contingent Convertible Securities: Evidence from the Global Banking Industry. *European Financial Management*, 23(1):127–152.

Avdjiev, S., Bogdanova, B., Bolton, P., Jiang, W., and Kartasheva, A. (2020). CoCo Issuance and Bank Fragility. *Journal of Financial Economics*.

Avdjiev, S., Bolton, P., Jiang, W., Kartasheva, A., and Bogdanova, B. (2015). CoCo Bond Issuance and Bank Funding Costs. *BIS and Columbia University Working Paper*.

Avdjiev, S., Kartasheva, A., and Bogdanova, B. (2013). CoCos: A Primer. *BIS Quarterly Review*, pages 43–56.

Barth, M. E. and Landsman, W. R. (2010). How did Financial Reporting Contribute to the Financial Crisis? *European Accounting Review*, 19(3):399–423.

BCBS (2010). Basel III: A Global Regulatory Framework for more Resilient Banks and Banking Systems. Technical report.

BCBS (2015a). Consultative Documentation – Guidance on Accounting for expected Credit Losses.

BCBS (2015b). Discussion Paper – Regulatory Treatment of Accounting Provisions.

Beatty, A. and Liao, S. (2011). Do Delays in Expected Loss Recognition affect Banks' Willingness to lend? *Journal of Accounting and Economics*, 52(1):1–20.

Beatty, A. and Liao, S. (2014). Financial Accounting in the Banking Industry: A Review of the Empirical Literature. *Journal of Accounting and Economics*, 58:339–383.

Belsley, D. A. (1991). *Conditioning Diagnostics – Collinearity and Weak Data in Regression*. Wiley Series in Probability and Statistics. Wiley.

Benczur, P., Cannas, G., Cariboni, J., Di Girolamo, F., Maccaferri, S., and Giudici, M. P. (2017). Evaluating the Effectiveness of the new EU Bank Regulatory Framework: A Farewell to Bail-Out? *Journal of Financial Stability*, 33:207–223.

Benoit, S., Colliard, J.-E., Hurlin, C., and Pérignon, C. (2017). Where the Risks Lie: A Survey on Systemic Risk. *Review of Finance*, 21(1):109–152.

Berger, A. N. and Bouwman, C. H. (2013). How does Capital affect Bank Performance during Financial Crises? *Journal of Financial Economics*, 109(1):146 – 176.

Beutel, J., List, S., and von Schweinitz, G. (2019). Does Machine Learning help us predict Banking Crises? *Journal of Financial Stability*, 45.

Billio, M., Getmansky, M., W., L. A., and Pelizzon, L. (2012). Econometric Measures of Connectedness and Systemic Risk in the Finance and Insurance Sectors. *Journal of Financial Economics*, 104(3):535–559.

BIS (2006). Supervisory Review Process and Market Discipline-Part 3: The Second Pillar-Supervisory Review Process.

BIS (2018). Countercyclical capital buffer (CCyB). `https://www.bis.org/bcbs/ccyb/`.

Bischof, J., Laux, C., and Leuz, C. (2019). Accounting for Financial Stability: Lessons from the Financial Crisis and Future Challenges. Working Paper, European Corporate Governance Institute.

Bisias, D., Flood, M., Lo, A., and Valavanis, S. (2012). A Survey of Systemic Risk Analytics. *Annual Review of Financial Economics*, 4(1):255–296.

Black, F. and Cox, J. C. (1976). Valuing Corporate Securities: Some Effects of Bond Indenture Provisions. *The Journal of Finance*, 31(2):351–367.

BOE (2016). Stress Testing the UK Banking System: 2016 Guidance for participating Banks and Building Societies.

Boermans, M. A. and van Wijnbergen, S. (2018). Contingent Convertible Bonds: Who invests in European CoCos? *Applied Economics Letters*, 25(4):234–238.

Borio, C., Drehmann, M., and Tsatsaronis, K. (2012). Stress-testing Macro Stress Testing: does it live up to Expectations? BIS Working Papers 369, Bank for International Settlements.

Borio, C. and Lowe, P. (2001). To Provision or Not to Provision. BIS Quarterly Review, Bank for International Settlements.

Boyd, J. H. and Runkle, D. E. (1993). Size and Performance of Banking Firms: Testing the Predictions of Theory. *Journal of Monetary Economics*, 31(1):47 – 67.

Breitung, J. and Knüppel, M. (2018). How far can we forecast? Statistical tests of the predictive content. Discussion Papers 07/2018, Deutsche Bundesbank.

Brownlees, C. and Engle, R. F. (2016). SRISK: A Conditional Capital Shortfall Measure of Systemic Risk. *The Review of Financial Studies*, 30(1):48–79.

Brunnermeier, M. K. and Pedersen, L. H. (2008). Market Liquidity and Funding Liquidity. *The Review of Financial Studies*, 22(6):2201–2238.

Bushman, R. M. and Williams, C. D. (2015). Delayed Expected Loss Recognition and the Risk Profile of Banks. *Journal of Accounting Research*.

Cahn, A. and Kenadjian, P. (2014). Contingent Convertible Securities: from Theory to CRD IV. *Institute for Law and Finance Working Paper Series*, 1(143).

Calomiris, C. W. and Mason, J. R. (2003). Fundamentals, Panics, and Bank Distress During the Depression. *American Economic Review*, 93(5):1615–1647.

Cebula, R. (2010). Determinants of Bank Failures in the US revisited. *Applied Economics Letters*, 17(13):1313–1317.

Chan, S. and Van Wijnbergen, S. (2014). CoCos, Contagion and Systemic Risk. *Working Paper*.

Chan, S. and Van Wijnbergen, S. (2016). CoCo Design, Risk Shifting and Financial Fragility. *CEPR Discussion Paper No. DP11099*.

Chan-Lau, J. A. (2010). Regulatory Capital Charges for Too-Connected-to-Fail Institutions: A Practical Proposal. *IMF Working Paper*.

Chen, N., Liu, X., and Yao, D. D. (2016). An Optimization View of Financial Systemic Risk Modeling: Network Effect and Market Liquidity Effect. *Operations Research*, 64(5):1089–1108.

Coffee Jr., J. C. (2011). Systemic Risk after Dodd-Frank: Contingent Capital and the Need for Regulatory Strategies beyond Oversight. *Columbia Law Review*, 111:795.

Cole, R. A. and Gunther, J. W. (1995). Separating the Likelihood and Timing of Bank Failure. *Journal of Banking & Finance*, 19(6):1073–1089.

Cole, R. A. and Gunther, J. W. (1998). Predicting Bank Failures: A Comparison of On- and Off-Site Monitoring Systems. *Journal of Financial Services Research*, 13(2):103–117.

Cole, R. A. and White, L. J. (2012). Déjà Vu All Over Again: The Causes of U.S. Commercial Bank Failures *This* Time Around. *Journal of Financial Services Research*, 42:5–29.

Constâncio, V. (2016). The Role of Stress Testing in Supervision and Macroprudential Policy. *Stress Testing and Macroprudential Regulation*, pages 51–68.

Cont, R., Duffie, D., Glasserman, P., Rogers, C., and Vega-Redondo, F. (2016). Preface. In *Special Issue on Systemic Risk: Models and Mechanisms*. Institute for Operations Research and the Management Sciences. Operations Research.

Council of the European Union (2013). Council Regulation (EU) No 1024/2013 conferring specific tasks on the European Central Bank concerning policies relating to the prudential supervision of credit institutions.

Cox, R. and Wang, G. (2014). Predicting the US Bank Failure: A Discriminant Analysis. *Economic Analysis and Policy*, 44(2):202–211.

Daníelsson, J. (2019). *Global Financial Systems: Stability and Risk*. Pearson, 6 edition.

Daníelsson, J., Shin, H. S., and Zigrand, J.-P. (2013). Endogenous and Systemic Risk. In *Quantifying Systemic Risk*, pages 73–94. National Bureau of Economic Research.

de Bandt, O. and Hartmann, P. (2000). Systemic Risk: A Survey. Working Paper Series 35, European Central Bank.

De Jonghe, O. (2010). Back to the Basics in Banking? A micro-analysis of Banking System Stability. *Journal of Financial Intermediation*, 19(3):387 – 417.

de la Lastra, M. and Ramón, J. (2012). 2011 EBA Stress Testing in Europe. *Advanced Research in Scientific Areas*.

de Lis, S. F., Pagés, J. M., and Saurina, J. (2001). Credit Growth, Problem Loans and Credit Risk Provisioning in Spain. In *Marrying the macro- and micro-prudential Dimensions of Financial Stability*, volume 1, pages 331–353. Bank for International Settlements.

Deloitte (2013). Going up? The Impact of Impairment Proposals on Regulatory Capital.

Demirgüc-Kunt, A. (1989). Deposit-Institution Failures: A Review of Empirical Literature. *Economic Review*, 4:2–18.

Demirgüç-Kunt, A., Detragiache, E., and Merrouche, O. (2013). Bank Capital: Lessons from the Financial Crisis. *Journal of Money, Credit and Banking*, 45(6):1147–1164.

Demirgüç-Kunt, A., Detragiache, E., and Tressel, T. (2008). Banking on the Principles: Compliance with Basel Core Principles and Bank Soundness. *Journal of Financial Intermediation*, 17(4):511 – 542.

Demyanyk, Y. and Hasan, I. (2010). Financial Crises and Bank Failures: A Review of Prediction Methods. *Omega*, pages 315–324.

Deutsche Bank (2012). Financial Data Supplement 4Q2012. https://www.db.com/ir/en/download/FDS_4Q2012_15042013.pdf.

DeYoung, R. and Roland, K. P. (2001). Product Mix and Earnings Volatility at Commercial Banks: Evidence from a Degree of Total Leverage Model. *Journal of Financial Intermediation*, 10(1):54–84.

DeYoung, R. and Torna, G. (2013). Nontraditional Banking Activities and Bank Failures during the Financial Crisis. *Journal of Financial Intermediation*, 22:397–421.

Diamond, D. W. and Rajan, R. G. (2000). A Theory of Bank Capital. *The Journal of Finance*, 55(6):2431–2465.

Diamond, D. W. and Rajan, R. G. (2001). Liquidity Risk, Liquidity Creation, and Financial Fragility: A Theory of Banking. *Journal of Political Economy*, 109(2):287–327.

Döring, B. (2016). Systemic Risk Measures and their Viability for Banking Supervision. *Dissertation*.

Duffie, D. (2011). *How Big Banks Fail And What to do About It*. Princeton University Press.

Duffie, D. (2018). Financial Regulatory Reform After the Crisis: An Assessment. *Management Science*, 64(10):4835–4857.

Dugan, J. C. (2009). Loan Loss Provisioning and Pro-Cyclicality – Remarks by John C. Dugan, Comptroller of the Currency, before the Institute of International Bankers.

EBA (2016). Report on Results from the EBA Impact Assessment of IFRS 9.

EBA (2017). Report on Results from the second EBA Impact Assessment of IFRS 9.

EBA (2018a). 2018 EU-Wide Stress Test. https://eba.europa.eu/risk-analysis-and-data/eu-wide-stress-testing/2018.

EBA (2018b). 2018 EU-Wide Stress Test: Frequently Asked Questions.

EBA (2018c). 2018 EU-Wide Stress Test: Methodological Note.

EBA (2018d). First Observations on the Impact and Implementation of IFRS 9 by EU Institutions.

ECB (2010a). Analytical Models and Tools for the Identification and Assessment of Systemic Risks. *Financial Stability Review*.

ECB (2010b). New Quantitative Measures of Sbystemic Risk. *Financial Stability Review*.

ESRB (2017). Financial Stability Implications of IFRS 9. https://www.esrb.europa.eu/pub/pdf/reports/20170717_fin_stab_imp_IFRS_9.en.pdf.

ESRB (2018). Adverse macro-financial scenario for the 2018 EU-wide banking sector stress test, European System of Financial Supervision. https://www.esrb.europa.eu/mppa/stress/shared/pdf/eu-wide_stress_test_adverse_macrofinancial_scenario.pdf.

ESRB (2019). Announced CCyB Rates. https://www.esrb.europa.eu/national_policy/ccb/all_rates/html/index.en.html.

Estrella, A., Park, S., and Peristiani, S. (2000). Capital Ratios as Predictors of Bank Failure. *Economic Policy Review*, 6(2).

Fajardo, J. and Mendes, L. (2018). CoCos Bond and Systemic Risk. *Working Paper*.

Financial Stability Forum (2009). Report of the Financial Stability Forum on Addressing Procyclicality in the Financial System.

Fiordelisi, F., Pennacchi, G., and Ricci, O. (2019). Are Contingent Convertibles going-concern Capital? *Journal of Financial Intermediation*, page 100822.

Flannery, M. (2005). No Pain, No Gain? Effecting Market Discipline via reverse Convertible Debentures. *Capital Adequacy beyond Basel: Banking, Securities, and Insurance*, pages 171–196.

Flannery, M. (2014). Contingent Capital Instruments for Large Financial Institutions: A Review of the Literature. *Annual Review of Financial Economics*, 6(1):225–240.

Flannery, M. (2016). Stabilizing Large Financial Institutions with Contingent Capital Certificates. *Quarterly Journal of Finance*, 06(02):1650006.

Flannery, M., Hirtle, B., and Kovner, A. (2017). Evaluating the Information in the Federal Reserve Stress Tests. *Journal of Financial Intermediation*, 29:1–18.

Foglia, A. (2009). Stress Testing Credit Risk: A Survey of Authorities' Approaches. *International Journal of Central Banking*, 5(3):9–45.

Friedman, M. and Schwartz, A. J. (1963). *A Monetary History of the United States, 1867 - 1960*. Princeton University Press.

FSB, IMF, and BIS (2009). Guidance to Assess the Systemic Importance of Financial Institutions, Markets and Instruments: Initial Considerations. *Report to the G20 Finance Ministers and Central Bank Governors*.

G20 (2009). Declaration on Strengthening the Financial System – London Summit. Declaration.

Gadanecz, B. and Jayaram, K. (2009). Measures of Financial Stability - A Review. In *Proceedings of the IFC Conference on Measuring Financial Innovation and its Impact*, volume 31, pages 365–380. Bank for International Settlements.

Gebhardt, G. (2016). Impairments of Greek Government Bonds under IAS 39 and IFRS 9: A Case Study. *Accounting in Europe*, 13(2):169–196.

Giesecke, K. and Kim, B. (2011). Systemic Risk: What Defaults Are Telling Us. *Management Science*, 57(8):1387–1405.

Goetz, M. R. (2018). Competition and Bank Stability. *Journal of Financial Intermediation*, 35:57 – 69.

Guerry, N. and Wallmeier, M. (2017). Valuation of Diversified Banks: New Evidence. *Journal of Banking & Finance*, 80(C):203–214.

Gupta, A., Lu, Y., and Wang, R. (2018). Addressing Systemic Risk Using Contingent Convertible Debt. A Network Analysis. *Working Paper*.

Hakkio, C. S. and Keeton, W. R. (2009). Financial Stress: What Is It, How Can It Be Measured, and Why Does It Matter? *Kansas City FED Economic Review*.

Haldane, A. and Madouros, V. (2012). The Dog and the Frisbee. *Proceedings - Economic Policy Symposium - Jackson Hole*, pages 109–159.

Hanson, S. G., Kashyap, A. K., and Stein, J. C. (2011). A Macroprudential Approach to Financial Regulation. *Journal of Economic Perspectives*, 25(1):3–28.

Hart, O. and Zingales, L. (2011). A New Capital Regulation for Large Financial Institutions. *American Law and Economics Review*, 13(2):453–490.

Hartmann-Wendels, T., Pfingsten, A., and Weber, M. (2019). *Bankbetriebslehre*. Springer, 7 edition.

Haselmann, R. and Wahrenburg, M. (2018). How demanding and consistent is the 2018 Stress Test Design in Comparison to previous Exercises? Technical report, European Parliament.

Hashim, N., Li, W., and O'Hanlon, J. (2016). Expected-loss-based accounting for Impairment of Financial Instruments: The FASB and IASB proposals 2009 – 2016. *Accounting in Europe*, 13(2):229–267.

Hashim, N., O'Hanlon, J., and Li, W. (2015). Expected-loss-based Accounting for the Impairment of Financial Instruments: The FASB and IASB IFRS 9 Approaches.

Hoelscher, D. S. and Quintyn, M. (2003). Managing Systemic Banking Crises. *IMF Occasional Paper 224*.

Hoenig, T. (2014). Release of Fourth Quarter 2014 Global Capital Index.

Homar, T. and van Wijnbergen, S. J. (2017). Bank Recapitalization and Economic Recovery after Financial Crises. *Journal of Financial Intermediation*, 32:16 – 28.

Hoogervorst, H. (2014). Closing the Accounting Chapter of the Financial Crisis.

Houston, J. F., Lin, C., Lin, P., and Ma, Y. M. (2010). Creditor Rights, Information Sharing, and Bank Risk Taking. *Journal of Financial Economics*, 96(3):485 – 512.

Huang, X., Zhou, H., and Zhu, H. (2009). Assessing the Systemic Risk of a heterogeneous Portfolio of Banks during the recent Financial Crisis. Finance and Economics Discussion Series 2009-44, Board of Governors of the Federal Reserve System (U.S.).

Hugonnier, J. and Morellec, E. (2017). Bank Capital, Liquid Reserves, and Insolvency Risk. *Journal of Financial Economics*, 125(2):266–285.

IASB (2010). Variations of an Expected Loss Approach.

IASB (2013). Exposure Draft - Snapshot: Financial Instruments: Expected Credit Losses.

IASB (2014a). IFRS 9 Financial Instruments.

IASB (2014b). IFRS 9 Financial Instruments – Basis for Conclusions.

IASB (2014c). IFRS 9 Financial Instruments – Project Summary.

IASB (2018). Conceptual Framework for Financial Reporting 2018 – Basis for Conclusions.

Jungherr, J. (2018). Bank Opacity and Financial Crises. *Journal of Banking & Finance*, 97:157 – 176.

Kashyap, A. K., Rajan, R. G., and Stein, J. C. (2008). Rethinking Capital Regulation. *Maintaining Stability in a Changing Financial System*.

King, T. B., Nuxoll, D. A., and Yeager, T. J. (2006). Are the Caueses of Bank Distress Changing? Can Researchers Keep Up? *Federal Reserve Bank of St. Louise Review*, 88(1):57–80.

Köhler, M. (2015). Which Banks are more risky? The Impact of Business Models on Bank Stability. *Journal of Financial Stability*, 16:195 – 212.

Kolari, J., Glennon, D., Shin, H., and Caputo, M. (2002). Predicting large US Commercial Bank Failures. *Journal of Economics & Business*, 54(4):361–387.

Korte, J. (2014). Why Closing Failed Banks Helps the Real Economy. In Szambelańczyk, J., editor, *Safe Bank*, pages 85 – 90. Bank Guarantee Fund.

Kothari, S. P. and Lester, R. (2012). The Role of Accounting in the Financial Crisis: Lessons for the Future. *Accounting Horizons*, 26(2):335–351.

Koziol, C. and Lawrenz, J. (2012). Contingent Convertibles. Solving or Seeding the next Banking Crisis? *Journal of Banking & Finance*, 36(1):90–104.

Krüger, S., Röesch, D., and Scheule, H. (2018). The Impact of Loan Loss Provisioning on Bank Capital Requirements. *Journal of Financial Stability*, 36:114 – 129.

Kund, A.-G. (2018). Can Systemic Risk Measures explain Bank Defaults? *Working Paper*.

Laeven, L. and Levine, R. (2007). Is there a Diversification Discount in Financial Conglomerates? *Journal of Financial Economics*, 85(2):331–367.

Laeven, L. and Levine, R. (2009). Bank Governance, Regulation and Risk Taking. *Journal of Financial Economics*, 93(2):259 – 275.

Laeven, L. and Valencia, F. (2018). Systemic Banking Crises Revisited. IMF Working Papers, International Monetary Fund.

Landini, S., Uberti, M., and Casellina, S. (2018). Credit Risk Migration Rates Modeling as open Systems: A micro-simulation Approach. *Communications in nonlinear Science & Numerical Simulation*, 58:147–166.

Liao, Q., Mehdian, S., and Rezvanian, R. (2017). An Examination of Investors' Reaction to the Announcement of CoCo Bonds Issuance: A global Outlook. *Finance Research Letters*, 22:58 – 65.

Lloyd, S. (2018). IFRS 9 and Equity Instruments.

Löffler, G. and Raupach, P. (2013). Robustness and Informativeness of Systemic Risk Measures. Discussion Papers, Deutsche Bundesbank.

Lucas, R. E. L. (1976). Econometric Policy Evaluation: A Critique. *Carnegie-Rochester Conference Series on Public Policy*, 1:19 – 46.

MacKinlay, A. C. (1997). Event Studies in Economics and Finance. *Journal of Economic Literature*, 35(1):13–39.

Maes, S. and Schoutens, W. (2012). Contingent Capital: An In-Depth Discussion. *Economic Notes*, 41(1-2):59–79.

Mahadeva, L. and Robinson, P. (2004). *Unit Root Testing in a Central Bank.* Centre for Central Banking Studies, Bank of England.

Martin, D. (1977). Early Warning of Bank Failure – A logit Regression Approach. *Journal of Banking & Finance*, 1:249–276.

Marton, J. and Runesson, E. (2017). The predictive Ability of Loan Loss Provisions in Banks – Effects of Accounting Standards, Enforcement and Incentives. *The British Accounting Review*, 49(2):162–180.

McDonald, R. L. (2013). Contingent Capital with a dual Price Trigger. *Journal of Financial Stability*, 9(2):230–241.

Mundlak, Y. (1978). On the Pooling of Time Series and Cross Section Data. *Econometrica*, 46(1):69–85.

Novotny-Farkas, Z. (2016). The Interaction of the IFRS 9 Expected Loss Approach with supervisory Rules and Implications for Financial Stability. *Accounting in Europe*, 13(2):197–227.

Pagano, M., Langfield, S., Acharya, V. V., Boot, A., Brunnermeier, M. K., Buch, C., Hellwig, M. F., Sapir, A., and van den Burg, I. (2014). Is Europe Overbanked? Report of the advisory scientific committee, European Systemic Risk Board.

Paisley, J. (2017). Stress Testing: Where next? *Journal of Risk Management in Financial Institutions*, 10(3):224–237.

Pankoke, D. (2014). Sophisticated vs. Simple Systemic Risk Measures. Working Papers on Finance 1422, University of St. Gallen, School of Finance.

Patro, D. K., Qi, M., and Sun, X. (2013). A simple Indicator of Systemic Risk. *Journal of Financial Stability*, 9(1):105–116.

Radjan, U., Seru, A., and Vig, V. (2010). Statistical Default Models and Incentives. *American Economic Review*, 100(2):506–510.

Raviv, A. (2004). Bank Stability and Market Discipline: Debt-for-Equity Swap versus Subordinated Notes. *Working Paper*.

Reinhart, C. M. and Rogoff, K. S. (2014). Recovery from Financial Crises: Evidence from 100 Episodes. *American Economic Review*, 104(5):50–55.

Reitgruber, W. et al. (2015). Methodological Thoughts on Expected Loss Estimates for IFRS 9 impairment: Hidden Reserves, Cyclical Loss Predictions and LGD Backtesting. Technical report, Erste Group Bank AG.

Riebl, L. and Gutierrez, P. (2018). A Review of Stress Test Methodology. *Journal of Securities Operations & Custody*, 10(3):254–267.

Roy, A. D. (1952). Safety First and the Holding of Assets. *Econometrica*, 20(3):431–449.

Sahin, C., de Haan, J., and Neretina, E. (2020). Banking Stress Test Effects on Returns and Risks. *Journal of Banking & Finance*, 117:105843.

Sánchez Serrano, A. (2018). Financial Stability Consequences of the Expected Credit Loss Model in IFRS 9. *Revista de Estabilidad Financiera*, 34:83–99.

Schuermann, T. (2014). Stress Testing Banks. *International Journal of Forecasting*, 30(3):717–728.

Schularick, M., Steffen, S., and Tröger, T. (2020). Bank Capital and the European Recovery from the COVID-19 Crisis. *SAFE White Paper No. 69*.

Schwarcz, S. L. (2008). Systemic Risk. *Georgetown Law Journal*, 97:193–249.

Segoviano, M. A. and Goodhart, C. (2009). Banking Stability Measures. Working paper, International Monetary Fund.

Squam Lake Working Group (2009). An expedited Resolution Mechanism for distressed Financial Firms: Regulatory hybrid Securities. *Council on Foreign Relations*, 10.

Steffen, S. (2014). Robustness, Validity and Significance of the ECB's Asset Quality Review and Stress Test Exercise. Study for the European Parliament.

Tong, X. (2015). Modeling Banks' Probability of Default. *Applied Economics and Finance*, 2(2).

Tsesmelidakis, Z. and Merton, R. C. (2013). The value of implicit guarantees. *Working Paper*.

Tukey, J. W. (1977). *Exploratory Data Analysis*. Addison-Wesley.

Vallascas, F. and Keasey, K. (2012). Bank Resilience to Systemic Shocks and the Stability of Banking Systems: Small is beautiful. *Journal of International Money and Finance*, 31:1745 – 1776.

Vaněk, T., Hampel, D., et al. (2017). The Probability of Default Under IFRS 9: Multi-period Estimation and Macroeconomic Forecast. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*, 65(2):759–776.

Vazza, D. and Kraemer, N. W. (2018). Default, Transition, and Recovery: 2017 Annual Global Corporate Default Study And Rating Transitions. *S&P Global Ratings*.

Vyas, D. (2011). The Timeliness of Accounting Write-Downs by U.S. Financial Institutions During the Financial Crisis of 2007–2008. *Journal of Accounting Research*, 49(3):823–860.

Wagner, W. (2010). Diversification at financial institutions and systemic crises. *Journal of Financial Intermediation*, 19(3):373 – 386.

Weiß, G. N. F., Bostandzic, D., and Neumann, S. (2014). What Factors drive Systemic Risk during international Financial Crises? *Journal of Banking & Finance*, 41:78–96.

Wheelock, D. and Wilson, P. (2000). Why do Banks Disappear? The Determinants of U.S. Bank Failures and Acquisitions. *The Review of Economics and Statistics*, 82(1):127–138.

Zaghdoudi, T. (2013). Bank Failure Prediction with Logistic Regression. *International Journal of Economics and Financial Issues*, 3(2):537–543.