THREE ESSAYS ON FINANCIAL STABILITY

Inauguraldissertation zur Erlangung des akademischen Grades
eines Doktors der Wirtschaftswissenschaften des Fachbereichs
Wirtschaftswissenschaften der Universität Osnabrück

vorgelegt
von
Michael Abendschein

Osnabrück, September 2020

Dekanin:	Prof. Dr. Valeriya Dinger
Referenten:	Prof. Dr. Peter Grundke
	Prof. Dr. Valeriya Dinger
Tag der Disputation:	25.09.2020

Acknowledgements

This work is not the result of a one-man show, but it has profited tremendously from other people's input and advice. It has been a pleasure to work with such brilliant colleagues and co-authors to whom I am truly indebted and whose substantial contributions have greatly improved this work.

First and foremost, I would like to thank my supervisor and co-author Peter Grundke for his constant assistance, his detailed and well-informed suggestions and for giving me the opportunity to follow my own ideas. I would also like to express my gratitude to my second co-author Gibran Watfe for his excellent contributions and for his creativity in finding loopholes at various seemingly hopeless stages of our project. That is to say, this work has not only benefited from his admirable research skills, but in particular from his friendship. This holds true for my third co-author Harry Gölz, too. I am not only thankful for his obvious and indispensable input to the thesis, but also for our insightful discussions and his detailed explanations that helped me become a better economist.

Several other people have enriched the years at the university, even though I will not be able to mention all of them. On their behalf, I would particularly like to thank my former colleagues Dietmar Holle, André Kühn and Elisabeth Krajewski who made me feel home at the new university very quickly; later Kai Rohde and Michael Tuchscherer for sharing our wonderful office and, finally, Chris Merklein for keeping up the good life outside the office, too.

There is no doubt that writing a thesis is not always fun. What is fun, however, is having a couple of days off in a Mediterranean locality with mountains, lakes, bars and cafes. Providing and sharing these moments (and maybe also his amazing, sometimes desperately needed programming skills) convinced me that Kolja Junginger should not be missing from this list. His contributions are manifold and highly appreciated.

Of course, I am most grateful to my family for their wise and caring support, and to my love, Angelika, to whom I owe the overwhelming joy of leaving Osnabrück with so much more than I could have ever imagined.

Contents

List of tables

List of figures

List of abbreviations

API ..... *Application programming interface*
ARMA..... *Autoregressive moving average*
BCBS .....*Basel Committee on Banking Supervision*
CGI..... *Closing-the-gap index*
D-SIB .....*Domestic systemically important bank*
EMH..... *Efficient market hypothesis*
GDP.....*Gross domestic product*
GJR-GARCH-DCC …… *Glosten Jagannathan Runkle - generalized autoregressive conditional heteroscedasticity - dynamic conditional correaltions*
G-SIB .....*Global systemically important bank*
HAR .....*Heterogeneous autoregressive model*
HHI..... *Herfindahl-Hirschman-Index*
IFRA..... *International financial regulation agreement*
IMF.....*International Monetary Fund*
MES ..... *Marginal expected shortfall*
MSE ..... *Mean-squared error*
ROA .....*Return-on-assets*
RV .....*Realized volatility*
SIFI.....*Systemically important financial institution*
SRM ..... *Systemic risk measure*
USD..... *US-Dollar*
VaR .....*Value-at-Risk*
VAR ..... *Vector autoregression*
WDI Database ..... *World development indicators Database*

List of symbols

B……global benefits

c……cost parameter

C……cost function

D……domestic benefits

F……empirical distribution of standard normally distributed random variable

$\mathcal{F}$……information set

h……benefit parameter

H……variance-covariance matrix

J……number of intraday intervals

k……regulatory capital ratio

L…… liabilities

n……number of cooperating countries

N……total number of countries

p……stock price

q……supervisory effort

r……stock/equity return

R……rank of systemic risk measure

s……seasonal standardization factor

τ……threshold parameter

U……(price) jump

w……weighting factor

W_t……Brownian motion

Y_t……count process

Z……standard normally distributed random variable

α……public good parameter

β……regression coefficient

γ……standardized correlation

δ……fixed effects

ε……error term

η……size of (price) jump

H……number of equidistant intervals within one day

θ……competition parameter

κ…….Gaussian kernel function

K……..…….integral over Gaussian kernel function

λ……..……bandwidth parameter

μ……..……..…….drift parameter

v……...…… error term

ξ…….....……error term

π…….....……. pay-off

ρ……...……correlation

ψ……...……probability/confidence level

σ……...……standard deviation

1 General introduction and overview

The global financial crisis has set the tone within a large fraction of the economic and financial literature for over a decade. The analysis of causes and consequences of this in part unprecedented financial crash has been of utmost importance not only for academics, but also for supervisors, policy makers and practitioners within the industry. Still, ongoing research keeps contributing to completing the picture that we have regarding what has happened from 2007 to 2009. Naturally, early efforts have primarily focused on a general understanding; see, for instance, Brunnermeier (2009) for an excellent contribution. Building upon these findings, consequences for reforms aiming at a partial redesign and stabilization of the global financial system are undertaken. This put financial regulators and supervisors at the forefront of ongoing policy debates. Former regulations that are suspected of having contributed to the crisis, as for instance put forward in Acharya et al. (2013), are constantly replaced and developed further. The Basel Committee on Banking Supervision (BCBS) as one of the most important international agencies dealing with questions with respect to global financial stability constantly reviews regulations in place and proposes reforms and amendments. The Basel reforms on behalf of the so-called Basel III framework were in part a reaction to misleading regulations before the crisis. One lesson incorporated in this post-crisis regulatory framework was the recognition that banks possess risk potential over and above the simple microprudential risk as measured for example by the bank-individual Value-at-Risk. The failure of Lehman Brothers and the resulting shockwaves perceptible all across the globe have unambiguously highlighted macroprudential risks within the banking system that have been overseen so far. The so-called systemic risk (not to be confused with systematic risk) has been quickly found to explain at least partly how the failure of a single bank could translate into nearly apocalyptic conditions on a global scale. Systemic risk is understood as the risk that the whole banking system collectively suffers large losses, which are then spread further towards other banks and economies, see Benoit et al. (2017). Verbally, this is an intuitive definition but is quickly turns out that it is very challenging to measure the systemic risk of banks concretely. The BCBS acknowledges five factors that jointly contribute to a bank's systemic risk potential, namely its size, interconnectedness, substitutability, complexity and cross-jurisdictional activity, see BCBS (2013). An indicator approach is helpful to get an intuition of the underlying drivers of systemic risk, but is does not allow for an objective, theory-based measurement thereof. Further market-based measures are developed that aim at assigning systemic risk potential to individual banks or the financial system as a whole that rely on market data such as equity returns or bank data such as

leverage. Even though the reasoning behind the respective computation is comprehensible, it turns out that different measures yield considerably different results regarding the assignment of systemic risk potential to specific banks, see e.g. Benoit et al. (2017) or Nucera et al. (2016).

This is the starting point of the first essay of this thesis that is co-authored with Peter Grundke. We empirically study the diverging results when ranking banks according to their individually assigned systemic risk ranks based on different systemic risk measures. Our objective is to understand which conditions and circumstances lead to more similar or more divergent systemic risk ranks such that the interpretation of such ranks can be facilitated and improved.

Even though macroprudential risks have gained momentum in the attention of academics and practitioners, there are of course several further facets of risks in financial markets that must not be neglected. One of the most prominent proxies for risk is volatility, or in the context of this work interchangeably called the standard deviation or variance of the return of a financial asset such as a stock. It is heavily studied for decades not only because it is of considerable importance for market participants, but also because it could be shown that volatility is to a certain degree predictable, see Poon and Granger (2003) for an excellent overview. The framework of this thesis does not allow for a comprehensive assessment of the plethora of approaches and techniques to measure, estimate, analyze and forecast volatility. Instead, the quite recent rise of social media as a platform (or, to be more precise, various platforms) to deliver, gain or change any kind of information on a global scale by a global audience offers new opportunities to study the interconnection between social media information and risks on financial markets. Once in a while, single Tweets, the name of messages posted on the social media platform Twitter, have shown to be able to destroy considerable amounts of stock market value within a very short time span of a few minutes. Irrespective of the potential truth or falsity of information provided in those Tweets, markets showed heavy reactions when for example an explosion at the White House has wrongly been announced in April 2013 which was accompanied by a market value loss of the S&P500 amounting to several billion US-Dollar, see e.g. Karppi and Crawford (2016). From a financial stability perspective, it is imperative to understand better the connections between social media information and stock volatility in order to be prepared for potential stabilizing actions. In the second essay of this work, which is co-authored with Gibran Watfe, we assess the usefulness of Twitter information in order to increase the forecast accuracy of stock return volatility. There is no consensus in the academic literature to what extent Twitter information can indeed be employed to improve volatility forecasts. This motivates us to study a comprehensively large fraction of the US stock market as well as the set of systemically important banks and some major stock indices in order to

determine whether Twitter information can yield insights regarding future volatility that goes beyond the forecasting power of traditionally employed models.

In the third essay, which is co-authored with Harry Gölz, we go one step back and do no longer consider specific regulations or specific transmission channels of shocks as given, but investigate why and under which circumstances financial regulators are motivated and incentivized to set-up common regulatory standards such as the Basel framework mentioned above. While it might appear intuitive that financial supervisors in a financially interconnected world engage in cooperative agreements, it is neither obvious that they do cooperate nor that they do so sufficiently. When common supervisory frameworks are designed and negotiated, tough discussions lead the way to any final agreement that is marked by opposing incentives on an individual country level, see for instance Goodhart (2011) with a history of the Basel accords. A typical trade-off represents the implementation of additional regulations yielding decreasing bank profits, with a different impact within the involved countries, of course. It is the remit of economists to understand the drivers of regulatory cooperation and to assess the efficiency of potential cooperative outcomes. Similar cooperative situations are heavily studied for instance with respect to trade agreements, environmental economics or industrial organization and there is reason to believe that the same rational might deliver valuable insights for regulatory cooperation as well.

In the following, the three essays are shortly summarized. Note that this is only for introductory purposes. Further details can be found in the respective chapters in the text.

For Essay 1, we compute systemic risk ranks for a large sample of international financial institutions. We follow popular approaches in the literature and use three of the most widely used systemic risk measures (SRMs), namely the *Marginal Expected Shortfall* (MES, see Acharya et al. (2017)), SRISK (Brownlees and Engle (2017)) and *exposure ΔCoVaR (*Adrian and Brunnermeier (2016)) that are computed based on publicly available data such as equity returns and leverage. In addition, we employ several modifications of these basic measures. Subsequently, we compute ranks based on all of these SRMs for each bank and each day. For our benchmark setting, we then compute quarterly bank-individual rank correlations for each rank pair. They form the heart of our analysis. We can see that correlations vary considerably across our sample. The ranks based on some SRMs, such as MES and ΔCoVaR, seem to be quite closely related while others turn out to produce a quite different systemic risk assessment. In the main part of the paper, we investigate potential drivers of these differences by conducting panel regressions where we test for the influence of various bank-individual and macroeconomic variables on rank correlations. Our results

indicate that only few variables are statistically and economically significantly associated with rank correlations. Macroeconomic variables seem to be more strongly associated with rank correlation, such as the unemployment rate or the Herfindahl-Hirshmann-Index as a proxy for market concentration in the financial industry. Moreover, the relation between rank correlation and the tested set of variables seems to be depending on the overall financial environment. Rank correlations are seemingly unaffected by most variables during high-volatility market phases while several variables are found to have a measureable influence during low-volatility phases. The overall results highlight the necessity to evaluate outcomes of systemic risk measures carefully. Loosely speaking, the more the better, i.e. it seems to be advisable to assess systemic risk potential of banks based on various measures simultaneously instead of only one or few single measures.

In Essay 2, we construct a sample of approximately 150 stocks that is meant to represent an important fraction of the stock market. Therefore, we include the constituents of the S&P100 index as well as the set of global systemically important banks and some of the largest stock indices. We do not want to find a single stock that can be forecasted most precisely with the help of Twitter data, but we want to conclude more generally whether or not and to what extent the precision of volatility forecasts can be improved by including Twitter information. For each single stock return, we compute volatility forecasts based on several volatility proxies and based on two of the most commonly used volatility models, that is the *heterogeneous autoregressive model* (HAR, see Corsi (2009)) and the HAR model enhanced by a jump component (see Andersen et al. (2007)). Twitter information are incorporated by counting the number of Tweets within a given time span that can be related to the specific stock under investigation. We follow two different approaches by counting, on the one hand, Tweets that contain so-called cashtags, i.e. the dollar sign $ plus the ticker symbol of a stock, a common procedure to signal that a Tweet relates to financial information regarding a specific stock. On the other hand, we count Tweets that contain the name of a respective company. We determine the in-sample fit of various models and in particular compare the out-of-sample forecasting performance of the various model specifications. Our results reveal that on average Twitter enhanced models do not show an increased forecasting precision. There are few stocks whose volatility can indeed be forecasted more precisely, but in general there is no superior performance of neither of the two approaches to count Tweets. Our overall results therefore suggest that the incorporation of Twitter information based on our straightforward implementation does not pay off, even though we cannot exclude that this might be valuable in some cases.

With Essay 3, we turn away from an empirical setting and start analyzing a theoretical model. The starting point of our analysis is the public good characteristic of financial regulation. Efforts to

regulate financial markets on a national level translate into an overall increase of financial stability and, hence, can lead to utility improvements for other countries as well. These positive externalities, however, are per se not incorporated into the utility maximization process of a national regulator or policy maker. This scenario prepares the ground for cooperation, because within a joint utility maximization these positive externalities are taken into account and a Pareto-improvement might become feasible. This scenario creates drawbacks in form of incentives to free-ride. On the national level, it is worthwhile not to contribute to financial stability when all the others contribute, given that costs of regulation have to be borne individually. In our model, costs materialize in form of reduced profitability of the national banking sector. The net benefit function of each regulator is assumed to be positively affected by efforts to maintain financial stability (or, in other words, the provision of financial regulation) on the national as well on the global level and by the profitability of the national banking sector. We study a partial cooperative solution where some countries decide to cooperate, based on an approach applied first in the literature on environmental agreements, see Finus and Rübbelke (2013). This approach allows determining the number of coalition members endogenously. A coalition consisting of two countries is stable for the vast majority of relevant parameters; depending on the specific functional form of the utility function also coalitions of three countries might be stable. Larger agreements, however, cannot be stable since incentives to leave the coalition are overwhelming. Still, we can show that exclusive club benefits that become effective within a coalition might increase the number of countries coordinating their supervisory efforts. Our model permits, on the one hand, to explain why it is hard to establish stable supervisory coalitions while it simultaneously gives valuable insights on how to influence policies that are meant to increase global financial stability by strengthening common regulatory oversight.

This dissertation concludes with a summarizing chapter where the most important results achieved across the three essays are presented in condensed form. Implications thereof and further interesting aspects are also sketched.

Contributions to the first essay "The ranking consistency of global systemic risk measures: empirical evidence" by Michael Abendschein and Peter Grundke

The initial research proposal and research design has been developed by Peter Grundke. He also contributed to updates of the research design, the interpretation of the results and, to a minor extent, to the presentation of the results in the manuscript of the working paper.

Michael Abendschein solely carried out the data gathering and processing and all statistical analyses. He prepared most parts of the manuscript and also contributed to updates of the research design.

Michael Abendschein Peter Grundke

2 The ranking consistency of global systemic risk measures: empirical evidence (Michael Abendschein and Peter Grundke)

2.1 Introduction

The question on how systemic risk, i.e. the risk that many banks simultaneously suffer large losses and that these losses are then spread through the system (see Benoit et al. (2017)), and the systemic importance of single banks could be measured has been intensively discussed in the academic research in the past years. During the financial crisis of 2007-2009, it became obvious that a purely microprudential regulation of banks is not sufficient to ensure the stability of the financial system as a whole, but that this method has to be supplemented by a macroprudential approach. Since then, global systemic risk measures (SRM) have been developed, on the one hand, to identify those financial institutions whose collapse would have the most harmful effect on the financial system, and, on the other hand, to identify those financial institutions who would be most significantly affected by a financial distress on the system level.

Such SRMs could enable regulators to classify banks according to their systemic risk potential (for instance, to impose stricter capital requirements for these banks) or help to estimate and assess the aggregate systemic risk level in an economy. However, up to now, a scenario where banks are regulated based on their individual systemic risk measure is far from being realistic since all the developed SRMs deliver strongly divergent results when classifying banks as systemically risky or not (see, e.g., Benoit et al. (2013), Grundke (2019), Nucera et al. (2016)). It is of particular interest to regulators and the financial industry as a whole to understand the drivers behind the outcomes of different systemic risk measures and the rankings based on them. At first glance, it seems to be desirable that SRMs assign very similar systemic risk rankings to individual banks but in this work it is shown that the correlation between the ranks based on popular systemic risk measures are quite volatile with values between 0.3 and 0.9 on average. However, as systemic risk is not a fully clear-cut concept, this might simply express the different facets of systemic risk that are measured by individual SRMs.

Given that several risk measures have already proved to contain valuable information (see Benoit et al. (2017) for an overview), it is not advisable to focus on one single metric. Still, it is hard to draw conclusions from systemic risk measures that inconsistently assign systemic risk rankings.

This work aims to shed light on the puzzle of inconsistent systemic risk rankings by investigating drivers of rank correlations based on popular systemic risk measures. Being able to explain potential factors that influence the rank correlation could allow for a more stringent and more accurate

application of SRMs in practice that could lead to a more efficient regulation. Given that particular bank characteristics or macroeconomic variables could have a pronounced effect on rank correlations, the outcome of systemic risk measures and especially the rankings thereof might be interpreted more reliably.

The results of a baseline panel regression with a sample of 80 large international banks indicate, however, that the relation between SRM rank correlations and a typical set of bank-individual and macroeconomic variables is rather unstable. Only few specific drivers can be detected that might explain differences in ranking outcomes. These results are highlighted and further inspected with the help of several robustness checks. We can show that the aforementioned unsatisfactory finding can be partly attributed to the fact that the relation between rank correlations and bank- and macroeconomic variables is strongly depending on the overall financial environment. Rank correlations are mostly unaffected by the considered variables during high-volatility market phases while several different variables are significantly associated with them during low-volatility phases. It turns out that rank correlations in general are more sensitive towards macroeconomic factors such as the unemployment rate, and to a minor degree towards factors that can be interpreted in a broader sense as proxies for the stability of a bank such as the market-to-book ratio and the loans-to-deposits ratio. In addition to the bank-level analysis, rank correlations are further investigated on an aggregate level, that is, we shift the focus from a purely bank-individual viewpoint and investigate drivers of correlations across various banks in a cross-sectional perspective.

This work confirms theoretical and empirical findings that show that ranks based on different systemic risk measures vary considerably (see, e.g., Benoit et al. (2013), (2017)) and, hence, that rank correlations are often not too large. Still, parts of the variation of rank correlations can be explained by bank-level variables and macroeconomic characteristics. However, given the considerable amount of remaining variation in our data, we suspect the underlying processes that lead to different systemic risk ranks to be not fully satisfactorily captured.

The paper is structured as follows: The literature on systemic risk measures is shortly sketched in Chapter 2.2. Chapter 2.3 explains methodological foundations with respect to systemic risk measures, rank computations and the baseline regression specification. Chapter 2.4 presents and summarizes the data employed in the study. In Chapter 2.5 we discuss the results of the panel regressions that include the baseline regression as well as a subsample analysis and an analysis based on weekly SRM data. Chapter 2.6 deals with the analysis of rank correlations on the macro-level including several additional subsample analyses. Chapter 2.7 discusses the achieved results

while Chapter 2.8 summarizes and concludes.

2.2 Literature

Up to now, regulatory authorities use an indicator-based approach to classify banks according to their systemic importance. The indicators are related to five broad categories: size, interconnectedness, lack of readily available substitutes or financial institution infrastructure, global (cross-jurisdictional) activity and complexity (see BCBS (2013)). In contrast to this, many other seemingly more sophisticated global SRMs relying on market and/or accounting data have been developed.

With respect to the literature on SRMs, various strands can be identified (for the following, see Grundke (2019)). First, there are papers in which the SRMs are introduced, motivated and empirically estimated the first time. Examples are Acharya et al. (2012), Acharya et al. (2017), Adrian and Brunnermeier (2016) or Brownlees and Engle (2016). Acharya et al. (2017) introduce the Marginal Expected Shortfall (MES) which measures how much an individual bank contributes to the overall risk of the banking system. Using bank-equity return data, MES is defined as the expected equity return of a bank conditional on the market return being smaller than some low quantile (e.g., 95%-quantile). The ΔCoVaR measure of a bank, proposed by Adrian and Brunnermeier (2016), corresponds to the increase of the conditional Value-at-Risk (VaR) of bank *i*, given that the market return of the whole banking system switches from its median to values at (below) some low quantile. This value is called the "exposure ΔCoVaR" of bank *i* by Adrian and Brunnermeier (2016); it measures the extent to which an individual bank is affected by systemic financial distress. To measure the influence that financial stress of bank *i* has on the whole banking system, Adrian and Brunnermeier (2016) also define the ΔCoVaR of the system where the conditioning is reversed ("contribution" ΔCoVaR). The ΔCoVaR of the system (with respect to bank *i*) denotes the increase of the VaR of the banking system conditional on bank *i* being in financial distress. The SRISK-index proposed by Acharya et al. (2012) and Brownlees and Engle (2016) is the expected capital shortfall of a bank conditional on a system crisis. The capital shortfall is understood as the capital reserves that a bank has to hold due to regulatory or prudential requirements minus its equity value. This SRM can be interpreted as an extension of the MES taking into account both the volume of the bank-liabilities and its size (measured by the bank's market capitalization).

Second, in a large body of follow-up papers, existing SRMs are taken (more or less) as they are and empirically estimated on various kinds of data sets. Partly, it is tried to identify determinants

that are significantly associated with the SRMs (e.g., features of the individual bank such as size or leverage or of the regulatory framework in which a bank operates). Examples of papers belonging to this second group are Döring et al. (2016), López-Espinosa et al. (2012) or Weiß et al. (2014a, 2014b). Döring et al. (2016) and Giglio et al. (2016) are also examples in which, the other way round, the prognostic power of SRMs for financial market and macroeconomic variables is checked. Buch et al. (2017) try to distinguish between contributions to systemic risk on the national and the European level.

Third, in another group of papers, modifications, extensions or enhanced strategies for the empirical estimation of existing SRMs are proposed. Examples are Girardi and Ergün (2013), Gravelle and Li (2013) or López-Espinosa et al. (2012, 2015). While, for example, Girardi and Ergün (2013) modify the conditioning event in the ΔCoVaR, López-Espinosa et al. (2012, 2015) modify it to capture asymmetric co-movements between system-wide and individual bank returns in case of a positive and a negative shock.

Finally, a forth group is formed by papers in which the ability of the proposed SRMs to consistently measure the systemic importance of financial institutions and the robustness of the SRMs are analyzed. Examples are Benoit et al. (2013), Benoit et al. (2017), Benoit et al. (2019), Danielsson et al. (2016a, 2016b), Grundke (2019), Grundke and Tuchscherer (2019), Jiang (2012), Löffler and Raupach (2018), and Nucera et al. (2016). Benoit et al. (2013) and Benoit et al. (2017) formally derive conditions that explain how the risk ranks based on MES, SRISK and ΔCoVaR are interrelated. Moreover, they show how the aforementioned SRMs can be computed with the help of systematic risk measures such as a firm's beta. This finding translates into interesting characteristics with respect to their joint rankings. While risk ranks of systemic and systematic risk measures can be identical in the cross-sectional perspective, i.e. for a given point of time across a sample of financial institutions, they can diverge in the time-series dimension for a given bank. Furthermore, Grundke (2019) investigates the rank consistency of different SRMs in a simulation-based banking network model and tests the ability of bank-specific variables as well as variables capturing network characteristics to explain rank consistency. Potential explanatory variables are found to be only weakly related to rank consistency.

In this paper, extend previous work with respect to the inconsistency of systemic risk ranks such as Nucera et al. (2016) by explicitly investigating potential factors that might be empirically associated with (in-)consistent risk ranks. For this, we rely on different specifications and various ro-

bustness checks. This procedure allows for a more detailed and systematic analysis of factors explaining diverging risk ranks compared to previous work in this field. While the relation between SRMs and several bank-specific as well as macroeconomic variables has already been established in previous work (e.g., Weiß et al. (2014a, 2014b) and Döring et al. (2016)), we amplify the analysis by testing if these variables are useful in explaining rank correlations. Our results allow, on the one hand, for a better understanding of the reliability of systemic risk ranks given different economic circumstances, which is crucial for policy making as well as scientific work. On the other hand, our results acknowledge and stress the opaqueness of the notion of systemic risk and highlight the necessity to take several different approaches to measure systemic risk into account when assessing the risk potential of the financial system and individual financial institutions, respectively.

2.3 Methodology

2.3.1 Systemic risk measures

Even though meanwhile there is a plethora of systemic risk measures available, in this study we focus on the seemingly most popular ones in the academic literature, namely Marginal Expected Shortfall (MES) and SRISK, and we use the (exposure) ΔCoVaR measure as well as modifications as explained in the following. It has to be noted that systemic risk in this work is solely measured with respect to the extent a bank is exposed to it, that is how much an individual bank suffers when the whole system experiences a stress period. On the contrary, systemic risk can also be addressed by measuring how much the financial stress of an individual institution affects the stability of the financial system as a whole. Thereby, the potential of a single institution to harm the financial system is attributed to systemic risk. Intuitively, ranks of two SRMs measuring, say, the exposure towards systemic stress, are more likely to be similar than ranks of two SRMs where the first one measures the exposure and the second one measures the contribution to systemic risk. In this work we only consider risk ranks of exposure SRMs, firstly because the most popular ones (MES, SRISK) belong to that group and secondly due to space constraints. More details with respect to the computation of the SRMs can be found in Technical Appendix.

Marginal Expected Shortfall (MES):

The MES was originally proposed by Acharya et al. (2017) and extensively estimated in various

empirical studies as the marginal contribution of a bank i to the overall systemic risk of an economy. In this study, a dynamic version of MES is used that is based on Brownlees and Engle (2016). In Acharya et al. (2017, p. 7), MES is defined as

$$MES_{i,t}(\tau) = -\frac{\partial ES_{m,t}(\tau)}{\partial w_{i,t}} = -E_{t-1}\left(r_{i,t} \middle| r_{m,t} < \tau\right) \tag{2.1}$$

where $ES_{m,t}$ is the expected shortfall of the market at time t given event τ defined as

$$ES_{m,t}(\tau) = E_{t-1}\left(r_{m,t} \middle| r_{m,t} < \tau\right) = \sum_{i=1}^{N} w_{i,t} E_{t-1}(r_{i,t} | r_{m,t} < \tau). \tag{2.2}$$

MES is multiplied with (-1) such that larger values represent larger systemic risk. $ES_{m,t}$ is the conditional expected market equity return r_m at time t given the market equity return is below some specified threshold τ. The market return is computed as the sum of the weighted individual equity returns $r_{i,t}$, based on the weights $w_{i,t}$ of the market capitalization of bank *i* in the economy at time *t*. The $MES_{i,t}(\tau)$ of bank i at time t is hence the marginal contribution of bank i to the overall expected shortfall of the market. Following Acharya et al. (2017) among others, τ is set equal to the 5%-quantile of the market return r_m. Acharya et al. (2017) mention that MES alone does not capture systemic risk effects that result from size and in particular leverage of specific firms. In order to obtain a measure that takes the size of a bank into account, Homar et al. (2016) (and in a similar fashion Adrian and Brunnermeier (2016) for ΔCoVaR) suggest computing a modification of MES by multiplying a bank's value for MES with its equity value. We follow this approach, too (MES-EQ in the following).

SRISK:

The SRISK-measure proposed by Acharya et al. (2012) and Brownlees and Engle (2016) is the expected capital shortfall of a bank i conditional on a systemic crisis. The capital shortfall is understood as the capital reserves that a bank has to hold due to regulatory or prudential requirements minus its equity value. SRISK is defined as (see Brownlees and Engle (2016, p.52)):

$$SRISK_{i,t} = k \cdot L_{i,t} - (1-k) \cdot (1 - LRMES_{i,t}) \cdot EQ_{i,t} \tag{2.3}$$

where $LRMES_{i,t}$ is the long-run MES of bank i at time t. The term "long-run" means that instead of daily equity returns, equity returns over some longer time horizon (e.g., 6 months) are employed. As a crisis scenario, Acharya et al. (2012) propose to consider situations in which the market drops by more than 40% over the next six months. Furthermore, k is the required prudential or regulatory

capital ratio. It is chosen at 8% since this is the standard in the academic literature, see for example Brownlees and Engle (2016). $L_{i,t}$ is the book value of total liabilities of bank i at time t and $EQ_{i,t}$ is the equity value of bank i at time t.

ΔCoVaR:

The third SRM is the (exposure-) ΔCoVaR measure as developed by Adrian and Brunnermeier (2016). The prefix "Co-" can be understood as "Conditional" since this SRM measures the Value-at-Risk of a bank's demeaned market equity return conditionally on the market being at its own Value-at-Risk $VaR_{m,t}(\psi)$. That is (Adrian and Brunnermeier (2016, p. 1710)):

$$P\left(r_{i,t} \leq CoVaR_t^{i|r_{m,t}=VaR_{m,t}(\psi)} \middle| r_{m,t} = VaR_{m,t}(\psi)\right) = \psi. \tag{2.4}$$

In order to compute $\Delta CoVaR_t$, the difference between the VaR of a bank given the market is in financial distress and the VaR of the bank given the market is not in financial distress is calculated. Financial distress of the market is defined as the market being at its VaR, whereas not being financially constrained is defined as the market being at its median state. That is (Adrian and Brunnermeier (2016, p. 1714)):

$$\Delta CoVaR_t(\psi) = -\left(CoVaR_t^{i|r_{m,t}=VaR_{m,t}(\psi)} - CoVaR_t^{i|r_{m,t}=VaR_{m,t}(0.5)}\right). \tag{2.5}$$

ΔCoVaR is multiplied with (-1) such that larger values represent larger systemic risk. Similar to MES, a modified measure of ΔCoVaR is computed by multiplying the ΔCoVaR-value of a specific bank with its equity value to account for size effects (ΔCoVaR-EQ).

The computation of MES, SRISK and ΔCoVaR follows exactly the procedure of Benoit et al. (2013).[1] They stress the well-known fact that there are several different ways to compute values for ΔCoVaR. While, originally, Adrian and Brunnermeier (2016) suggest using quantile regressions, another popular approach is to employ a GARCH-DCC model (see for example Benoit et al. (2013) and Döring et al. (2016)). The conditional volatility modelling allows measuring time-varying ΔCoVaR directly, while the quantile estimation approach requires time-varying macroeconomic (state) variables to estimate the conditional VaR. We follow Benoit et al. (2013) and

[1] Benoit et al. (2013) kindly provide their MATLAB code that they use to compute MES, SRISK and ΔCoVaR via their own open source project runmycode.org. Extensive technical details are provided in the appendix of their paper.

compute ΔCoVaR via a GJR-GARCH-DCC model, a procedure similarly used to compute MES and SRISK.

2.3.2 Rank correlations

To analyze the ranking consistency of the various SRMs, rank correlations are computed. For each day t and for each considered SRM $d \in \{1, \dots, D\}$, the rank $R_{d,i}^{t} \in \{1, \dots, N_t\}$ of bank i within the group of N_t banks that have survived until time t is computed. Larger values of the respective SRM at time t correspond to a larger systemic risk and, hence, lead to a higher rank ($R_{d,i}^{t} = N_t$: bank i has highest SRM d at time t, $R_{d,i}^{t} = 1$: bank i has smallest SRM d at time t). Figure 2.1 illustrates SRM ranks over time for five different SRMs for Goldman Sachs as an exemplary bank and it can be seen how strongly ranks vary for a single bank.

Afterwards, for each bank $i \in \{1, \dots, N_t\}$ and for each combination of two SRMs $d_1, d_2 \in \{1, \dots, D\}$, the Spearman correlation coefficient $Corr_{i,q}(R_{d_1,i}, R_{d_2,i})$ of the bank's ranks for a given quarter q is computed. This is what we call 'time-series rank correlation' in the following. Alternatively, in Chapter 2.6, a 'cross-sectional rank correlation' is computed.

The time window is chosen to contain all daily ranks during one quarter. That means the rank correlation $Corr_{i,q}\left(R_{d_1,i}, R_{d_2,i}\right)$ for quarter q is based on all daily SRM ranks during this quarter. This procedure is chosen in the baseline specification for two main reasons. First, with approximately 65 daily ranks per quarter, a sufficiently large number of ranks is considered to compute correlations. Second, since each correlation is based on non-overlapping data, that is, each rank is only used to compute one specific correlation, one reason for auto-correlated correlations is avoided.

Nucera et al. (2016) reject this time-series analysis since they assume (and confirm in their data based on a univariate Augmented-Dickey-Fuller test) that time series of ranks are frequently non-stationary. In contrast, applying the same procedure for our data, the hypothesis of a unit-root can be rejected for each SRM rank series for the vast majority of periods (not reported). This different finding might be explained by the different bank sample considered.

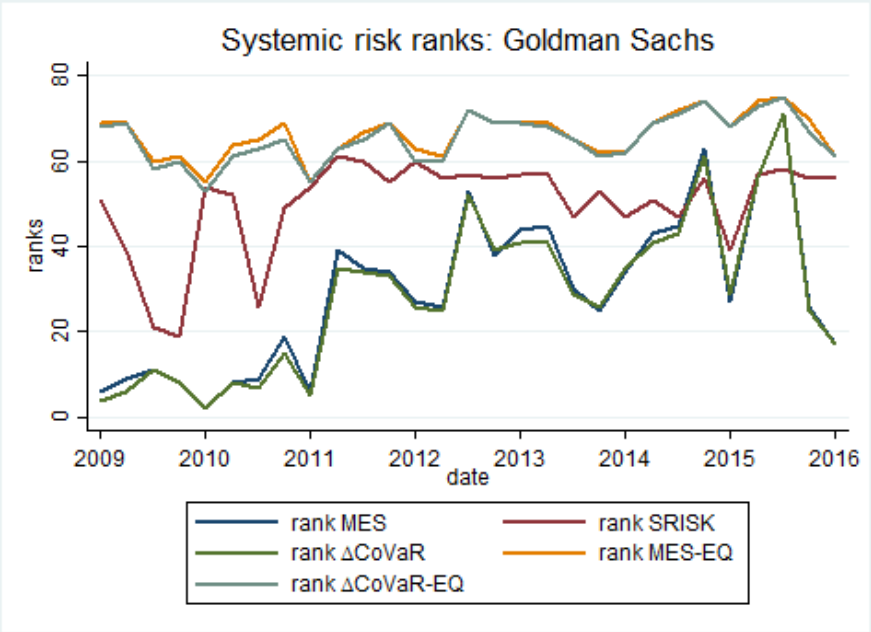


Figure 2.1: Systemic risk ranks

Systemic risk ranks of an exemplary bank, Goldman Sachs, for five different SRMs over the sample period from 2009 to 2016 are displayed. The higher the rank, the more systemic risk potential is assumed, given a sample of 80 financial institutions.

2.3.3 Baseline specification

For each pair $d_1, d_2 \in \{1, \dots, D\}$ of SRMs and based on quarterly data, panel regressions of the following type are run (with bank- and time-fixed effects to account for unobserved heterogeneity as well as standard errors clustered at the bank level):

$$Corr_{i,q}\left(R_{d_1,i}, R_{d_2,i}\right) = \beta_0 + \sum_{m=1}^{M} \beta_m^{d_1,d_2} \cdot Bank\ variables_{m,i,q-1} + \sum_{p=1}^{P} \beta_r^{d_1,d_2} \cdot Macroeconomic\ variables_{p,i,q} + \delta_i + \delta_q + \varepsilon_{i,q}^{d_1,d_2} \tag{2.6}$$

where $Corr_{i,q}\left(R_{d_1,i}, R_{d_2,i}\right)$ denotes the (time-series) correlation of bank i's rank in quarter q within the sample according to the SRMs d_1 and d_2.

The employed bank-specific and macroeconomic variables are described in the following chapter.

2.4 Data and hypotheses

We investigate a sample consisting of large American, Canadian as well as European banks. Many

related studies focus on US-American banks, in particular the early studies such as Acharya et al. (2017), Billio et al. (2012) or Brownlees and Engle (2016). American financial institutions are a natural starting point for the analysis due to their predominant role during the financial crisis. After establishing some risk measures, research focused on European banks as well, examples are Döring et al. (2016), Engle et al. (2015) or Nucera et al. (2016). While the crisis started in the US-financial system, it spread over to Europe soon thereafter illustrating the fragility of the European banking system and, hence, the need to understand and anticipate problems arising therein. As opposed to the US, the financial system in Europe is mainly marked by few very large banks per country (a notable exception is for example Germany, with large sectors of savings banks and cooperative banks). It is up to now doubtful whether or not those countries would be able (again) to save their banks in case of a systemic crisis. After all, we restrict our analysis to American, European and in addition Canadian banks. In addition, our sample selection is based on the asset size of banks, following i.e., Beltratti and Stulz (2012) or Döring et al. (2016). Banks are included if they are larger than 50bn in assets (in US-Dollar) at the end of 2008. In addition, all banks are included that are considered to be systemically important by the Basel Committee on Banking Supervision as of 2013 and national regulators, both on global and local level. However, some banks had to be excluded due to missing data. Subsidiaries are not included as long as their mother institution is part of the sample.[2] This exercise results in a sample of 80 banks, see Table A.1 in Appendix A.

Bank specific variables, mainly balance sheet items, are used at a quarterly frequency as they are reported from Bureau van Dijk's Bankscope database and cover a sample range from 2009 Quarter 3 to 2016 Quarter 3.

Macroeconomic variables are retrieved from World Development Indicators (WDI) Database, the International Monetary Fund (IMF) and national central banks. In this database, they are only available on an annual basis and, hence, they are transformed into quarterly data by using cubic splines. This is a standard technique in related work such as López-Espinosa et al. (2012) or Döring et al. (2016).

Market data, namely daily closing stock prices, daily closing price of the MSCI world index as well as daily data on market capitalization for the computation of the various systemic risk

[2] In case of a merger or acquisition, banks that meet the specified criteria are included as long as they are independent.

measures is retrieved from Thomson Reuters' Eikon database. Data on total liabilities that is required to compute SRISK stems from Bureau van Dijk's Bankscope as well.[3]

Potential explanatory variables are in particular chosen and included if they are found to have an impact in previous empirical work that studies drivers of levels of SRMs (see, e.g., Adrian and Brunnermeier (2016), Bostandzic et al. (2014), Buch et al. (2017), Cai et al. (2018), Girardi and Ergün (2013), Irresberger et al. (2017), López-Espinosa et al. (2012), Weiß et al. (2014a, 2014b), Zhou and Tarashev (2013)). The idea behind this specification is that if variations of these variables have a differential impact on the level of SRMs, they might also have a significant impact on the ranking consistency of SRMs. However, the aforementioned studies reveal very ambiguous results. It turns out that the relation between bank- or macro-variables on the one hand and the outcome of systemic risk measures on the other hand is very sensitive to the specific setting. While Döring et al. (2016), for instance, find a rather negative relation between leverage and various systemic risk measures using a VAR model, the results of Bostandzic et al. (2014) show in a standard regression setting that leverage and systemic risk of individual banks are rather positively connected. Similarly, inflation and GDP growth is found to be negatively associated with the degree of systemic risk of financial institutions measured by means of SRISK by Bostandzic et al. (2014) and Buch et al. (2017), while Döring et al. (2016) find a rather positive relation. A common finding in the literature is the rather weak explanatory power of most of the employed variables in terms of statistical significance. Hardly any bank characteristic or macroeconomic variable can explain the outcome of SRMs significantly across different specifications. Table A.2 (in Appendix A) lists a summary of linkages between bank- or macro-variables and systemic risk found in the literature.

Note, that there are no ex-ante hypotheses concerning the relation between the rank correlation of SRMs and the explanatory variables. The presented bank- and macro-variables, however, have been widely associated with the level of systemic risk in previous studies even though the expected direction of the influence is not always clear ex-ante. High leverage, for instance, could be assumed to be a driver of systemic risk given that it could lead to excessive risk taking, see, e.g. Hovakimian et al. (2012). On the contrary, higher leverage could be assumed to reduce a bank's risk seeking

[3] The MSCI World index is used as a proxy for the market return that is required for the computation of the various systemic risk measures. We follow the approach of Acharya et al. (2017) by considering a non-financial market index. Interestingly, the systemic risk measures turn out to be quite sensitive towards the choice of the respective market index highlighting the need to carefully assess and interpret SRM ranks.

by means of a more prudent management of its remaining liquid reserves, see Irresberger et al. (2017) and Vallascas and Hagendorff (2011). We embark on this approach and investigate the relation of the presented variables towards rank correlations of SRMs as well, without explicitly referring to an economic model that distinctly explains a potential channel. In this study and according to the related literature, the following bank-specific and country-specific macroeconomic variables are used as potential explanatory factors for the correlation of SRM ranks:

Assets: Assets are measured as the natural logarithm of total assets of a bank, a standard variable in the literature. It accounts for the size of the bank. A bank's size is one of the five indicators that the Basel Committee on Banking Supervision uses to define systemically important banks (see BCBS (2013)). Larger banks could take on more (systemic) risk since they could assume to be protected by implicit government guarantees (see, e.g., Bostandzic et al. (2014), Gandhi and Lustig (2015)).

Reserves-to-loans: This ratio measures potential losses that might occur due to a bank's bad performing credit portfolio. The higher the fraction of loss reserves compared to the overall amount of granted loans, the more negative might be the performance expectations of the bank (see e.g., Bostandzic et al. (2014)). Especially during economic downturns, this might be associated with increasing fragility of the bank.

Market-to-book: The market-to-book ratio is computed as the ratio of market valued equity to book valued equity and is often used in order to compare the market view (i.e., forward looking) with the valuation based on balance sheet data. Higher market-to-book ratios are associated with a lower system-wide risk potential (see Döring et al. (2016)).

Leverage: It is computed as the ratio of total debt to total assets (see, e.g., Weiß et al. (2014b)). It is associated with a higher default risk and a higher systemic risk (see e.g., Brunnermeier et al. (2012), Bostandzic et al. (2014)).

Return on assets (ROA): This is a typical profitability measure (see, e.g., Buch et al. (2017)). Its influence on financial stability is suspected to be ambiguous. While on the one hand, profitable banks can be expected to be more stable, on the other hand, excessive rent seeking can be a sign for a particularly risky business model as well.

Long-term funding: This variable measures the ratio of long-term funding to total funding. During the crisis, it has become apparent that a short-term funding structure, sometimes on a day-to-day basis, can pose considerable threat to the stability of a bank. Whenever roll-over funding becomes

difficult or expensive, banks with a larger share of short-term funding can get into trouble. Hence, larger values of the ratio of long-term funding to total funding are associated with a more stable capital structure because banks are less prone to liquidity risk (see, e.g., Brunnermeier and Pedersen (2009)).

Non-interest income: It is measured as the ratio of non-interest income to total interest income (see, e.g., Weiß et al. (2014a), Buch et al. (2017)). The lower the ratio, the stronger a bank's business model is supposed to be associated with the classical banking business, that is granting loans and receiving deposits which is assumed to be less risky in terms of systemic risk. In contrast, the higher the fraction of non-interest income, the stronger banks are assumed to be connected to global financial markets and, hence, more dependent and vulnerable to global financial stability concerns.

Loans-to-deposits: This ratio is typically used to account for the liquidity risk of a bank (see, e.g., Döring et al. (2016)). Higher ratios can be found for banks that do not mainly rely on deposits in order to finance their loans. Consequently, they are expected to face a higher liquidity risk in case of financial distress.

Tier1 ratio: This is a standard measure to account for the quality of the capital base. Higher Tier1 ratios are in a straightforward manner associated with more stable banks (see e.g., Bostandzic et al. (2014), Cihák et al. (2013)).

Z-Score: This risk measure for banks is also interpreted as a metric for a bank's distance to default. It is computed as the ratio of the sum of the equity-to-assets capital ratio and the return on assets and the standard deviation of the return on assets. A low value of the z-score of a bank is associated with a risky business model. It is also used as a proxy for systemic risk (see, e.g., Anginer et al. (2014), Li et al. (2017)).

Interconnectedness: This variable measures the extent a bank is interconnected with its peers by counting the number of Granger causal connections between each pair of bank equity returns (see Billio et al. (2012)). While Billio et al. (2012) find an increasing degree of interconnectedness during times of financial distress, Bostandzic et al. (2014) cannot establish a robust significant relation with respect to various SRMs.

As country-specific variables, *unemployment*, *inflation* and *GDP growth* are used. While the former is suspected to have a negative influence on financial stability, the latter two should positively affect financial and economic stability, at least in case of a moderate range of values (see e.g.,

Döring et al. (2016)).

Further macroeconomic indicators include the *Herfindahl-Hirschman–Index (*HHI*)* as a measure for market concentration based on a bank's asset size. Its influence on financial stability and systemic risk is ambiguous. While for instance Uhde and Heimeshoff (2009) find that markets that are more concentrated are more stable, Bostandzic et al. (2014) state an inverse relation between stability and market concentration with respect to several systemic risk measures.[4] Finally, a country's stock market conditions, that is, its *stock market return* and *volatility* as well as *long-term sovereign interest rates* are used as independent variables. All detailed definitions of the variables are given in Table A.3 in Appendix A.

Table A.4 (in Appendix A) gives an overview on summary statistics of the employed dependent and independent variables. Note that all explanatory variables are winsorized at the 1% and 99% level.[5] A larger set of additional variables has initially been available. The absence of multicollinearity, however, could not be confirmed in many cases. Thus, the final set of exogenous regressors is the result of eliminating any regressors that are found to induce multicollinearity.

It is interesting to see that all rank correlations contain values for approximately the whole range of possible outcomes from -1 to 1, illustrating the inconsistency of SRM ranks. The ranks of MES and exposure-ΔCoVaR seem to be more closely related than each of them with SRISK. While the mean correlation for the former pair is between 0.6 and 0.9 depending on whether or not 'pure' SRM or equity size-adjusted variants are considered, the mean correlations between SRISK and the other SRMs is around 0.3 to 0.4 only. Similarly, the standard deviations (across banks and over time) of rank correlations based on SRISK are larger than for other correlation pairs.

2.5 Results panel regressions

2.5.1 Baseline panel regressions

Results of the baseline regressions can be found in Table 2.1. It can be seen that the influence of bank- and macro-variables is ambiguous. Only some of the tested variables are significantly associated with the whole range of rank correlations, such as a bank's market-to-book ratio, its loans-

[4] Recently, the discussion with respect to the relation between market concentration and systemic risk has gained momentum. Beck et al. (2018) show that the unbundling of various facets of market concentration allows a more differentiated analysis of the different effects on systemic risk.

[5] Correlations are not winsorized as they are bounded by construction such that it remains rather unclear how outliers could be determined.

to-deposits ratio, and, on the macro-level, the unemployment rate as well as the HHI index as a measure for market concentration. In addition, bank size is partly found to be positive and significant for three out of ten correlations. This can be interpreted as a first indication that the systemic risk potential of larger banks is more consistently estimated across the different SRMs. The relation between rank correlations and a bank's market-to-book ratio and its loans-to-deposits ratio, respectively, appears to be even stronger. Four out of ten correlations are strongly significantly associated with the former one (with a positive sign) while five out of ten correlations are strongly significantly associated with the latter one (with a negative sign). Hence, the higher the market-to-book ratio, the more similar risk ranks turn out to be, and, in a similar fashion, the lower the loans-to-deposits ratio, the higher the SRM rank consistency. As higher market-to-book ratios are a sign of a positive outlook on future performance given that the forward-looking market valuation is relatively larger than the book valuation based on past performance, it could be interpreted that SRM ranks are more similarly estimated in case markets have a rather positive view on a bank's future business. In addition, Döring et al. (2016) note that higher market-to-book ratios are associated with lower systemic risk, such that our results indicate that SRM rank consistency might be higher when system-wide risk is lower, even though this finding cannot be confirmed in further robustness analyses. Next, loans-to-deposits is seen as a proxy for liquidity risk as a higher ratio means that relatively more loans are financed by more liquidity-sensitive instruments such as short-term borrowings that are more prone to sudden capital withdrawals compared to more stable deposits. Hence, the previous finding that assumes higher rank consistency for banks with a better outlook is somewhat strengthened by this result that indicates higher rank consistency for banks with a lower liquidity risk. The remaining bank variables are either insignificant or they exhibit a changing sign for different pairs of SRMs[6] that does not allow to draw further conclusions for the time being. Hence, even though there is weak evidence that more stable banks appear to exhibit more similar risk ranks, the overall picture reveals that it is difficult to assign larger explanatory power to a specific set of bank variables when it comes to explaining SRM rank correlations. This interpretation can be confirmed by looking at the economic significance of statistically significant coefficients. The largest coefficient for market-to-book ratio is estimated for the correlation between ranks of SRISK and MES-EQ with a value of 0.153. Taking a one-standard deviation increase of the market-to-book ratio into account (around 0.66), this could only lead to an increase of around 0.1 in rank correlations. Similarly, a one-standard deviation increase of the loans-to-

[6] This ambiguity is also found by Grundke (2019) in a simulation-based banking network model.

deposits ratio (also around 0.66) could at most lead to an increase in correlations by 0.07 for the rank correlation between MES and ΔCoVaR-EQ. In case a bank's asset size is statistically significant, its economic significance is slightly larger. A one standard deviation increase (1.23) could lead to an increase of up to 0.15 for the rank correlation between MES and MES-EQ and similarly to the rank correlation between MES-EQ and ΔCoVaR-EQ. In general, bank variables are not only weakly associated with rank correlations, the magnitude of economic significance is also limited in the few cases where statistical significance can be established.

With a view on macroeconomic variables, it can be seen that several ones are found to be statistically significantly associated with rank correlations. In particular, the unemployment rate (five out of ten correlations), the HHI (six out of ten) and the stock market return (four out of ten) appear to exhibit explanatory power. The remaining variables are once again either insignificant or the switching signs of the coefficients impede a straightforward interpretation. The HHI stands out compared to the other factors. It is not only found to be statistically significantly associated with rank correlations; the economic significance is also considerable. With a sample standard deviation of 0.05, a one standard deviation increase of the HHI can lead to an increase in correlations of up to 0.17 for the rank correlation between MES and ΔCoVaR-EQ, with similar values for further correlations. Hence, a lower market concentration in a country's financial industry is found to be quite strongly associated with higher SRM rank correlations. Even though unemployment and stock market returns are robustly significant in statistical terms, indicating higher rank correlations for higher unemployment and higher stock returns, the economic significance of these variables is not too pronounced. A one standard deviation increase in unemployment (0.05) could at most lead to an increase in rank correlations of around 0.06 between MES and MES-EQ, while similarly a one standard deviation increase of a country's stock market return (0.16) could only induce a maximum increase in correlations of around 0.02 for the correlation between ΔCoVaR and ΔCoVaR-EQ. These results suggest that macroeconomic variables do have a stronger effect than bank-individual factors; however, the influence appears to be rather diffuse. Macroeconomically negative developments such as rising unemployment rates and positive developments such as rising stock returns both seem to be associated with higher rank correlations.

When we compare the results for SRM pairs that exhibit a large average correlation with those with a smaller one, no clear trend can be depicted. The rank correlation between MES and ΔCoVaR as well as the correlation between MES-EQ and ΔCoVaR-EQ is very high with values around 0.9. The former pair of SRM rank correlations is rarely significantly associated with the tested set of variables (four out of eighteen) while, on the contrary, the latter pair is more often found to be

significantly associated (seven out of eighteen). A similar picture emerges when looking at SRM ranks that are rather weakly correlated. If at all, rank correlation based on SRISK seem to be more weakly associated with the tested set of variables with partly only three significant coefficients for regressions based on the correlation between SRISK and MES-EQ and ΔCoVaR-EQ, respectively, but the difference compared to the remaining correlations is not very large.

Until now, the baseline specification partly reveals that some factors might be indeed associated with rank correlations; however, the majority of variables is rather unrelated in the given setting. Note, that the (adjusted) R^2 is low throughout the different regressions. This is a typical result for panel data analyses with respect to systemic risk measures, see for instance Mühlnickel and Weiß (2015) for similar values.

2.5.2 Panel regressions for (non-) different volatile market phases

The results of the baseline panel analysis in Chapter 2.5.1 have already allowed gaining first insights on how SRM rank correlations can be associated with bank and macroeconomic variables. As a next step, we want to investigate whether or not these statistical relations differ during less and more volatile periods. The need to ascribe systemic risk potential and ranks to financial institutions is particularly relevant during times of financial distress when regulators need to identify the most vulnerable banks. Therefore, we use data from the CBOE Volatility Index (VIX), an index measuring the implied volatility of the S&P500, to rank the quarters in our sample based on its VIX value[7]. In a next step, the quarters of the upper 50% quantile are used as a subsample for rather distressed periods whereas the quarters of the lower 50% quantile are used as the subsample for stable periods. Given the comparably low number of quarters in our sample, a further reduction of, say, only the upper and lower 25% quantile would result in an unreasonable loss of data.

A first look at the correlations themselves illustrates that they tend to increase during the stress period. Both the mean and the median of the correlations are slightly larger for the more volatile subsample for the majority of SRM pairs, as can be seen in Table 2.2. This difference, however, is only found to be significant for some of the considered correlations.

[7] Even though the CBOE VIX is a measure of the volatility of the (US-American) S&P500 only, it is also understood as a general benchmark index for volatility on financial markets. For our sample period, the correlation between the VIX and the corresponding European EUROSTOXX50 Volatility index and the Canadian S&P/TSX60 Volatility Index is 0.9 and 0.85, respectively, so we advocate the use of the VIX to define (non-) volatile periods in our work.

The results of the subsample regression analyses are illustrated in Table A.5 and Table A.6 (in Appendix A) and a distinctly different picture emerges. First, it is obvious that rank correlations are more weakly associated with bank-explanatory variables when the more volatile subsample is considered. The market-to-book ratio is not found to be significant anymore; the asset size is still partly significant, however, with changing signs across the different specifications. The loans-to-deposits ratio is the only bank-specific variable that remains similarly significant, and the coefficients remain similar in size. A notable difference is the occurrence of the Tier 1 ratio in the set of significant variables. We can see that a considerable number of correlations is significantly affected by the Tier 1 ratio, but the signs of the coefficients are once again not stable. Overall, the association between bank-individual variables and SRM rank correlations tends to be much weaker during more volatile market phases times, with the exception of the loans-to-deposits ratio that is still negatively related to rank correlations as it is found in the baseline specification.

Results for the more volatile subsample turn out to be different with a view on macroeconomic variables. Unemployment and the HHI are still quite robustly significant across the various specifications. An increase of unemployment drives rank correlations upwards while a stronger market concentration drives correlations downwards. The size of the coefficients, however, is striking as it has merely doubled. While the standard deviation of the respective variables based on the more volatile subsample has not changed much (still around 0.05) [8], we can conclude that the economic significance of unemployment and HHI has also doubled. All rank correlations between MES and ΔCoVaR including a variant with equity size increase by around 0.3 in case of a one-standard deviation increase of HHI. Similarly, the same correlations decrease by around 0.14 following a one-standard deviation increase in unemployment.

It is very interesting to see that some rank correlations are particularly associated with a specific set of variables during more volatile periods. However, it is eye-catching that always the same set of correlations exhibits significant coefficients, namely those that include either MES or ΔCoVaR and one variant thereof that is based on the multiplication with equity size. Rank correlations including SRISK remain seemingly unaffected by the tested set of explanatory variables, so does the correlation between ranks of “pure” MES and “pure” ΔCoVaR. Thus, it might be suspected that the inclusion of equity size in the computation of SRMs may influence these results.

[8] Note that all standard deviations are computed based on the respective crisis or non-crisis sub-sample.

Table 2.1: Regression results – Baseline regression

	MES-MES-EQ	MES-SRISK	MES-ExpΔCo-VaR	MES-ExpΔCo-VaR-EQ	MES-EQ-SRISK	MES-EQ - ExpΔCo-VaR	MES-EQ-ExpΔCo-VaR-EQ	SRISK-ExpΔCo-VaR	SRISK-ExpΔCo-VaR EQ	ExpΔCo-VaR-ExpΔCo-VaR-EQ
Assets	0.121**	-0.0610	0.0059	0.0695	0.120	0.117**	0.0325**	-0.0473	0.106	0.0697
	(2.63)	(-0.69)	(0.79)	(1.48)	(1.18)	(2.52)	(2.52)	(-0.54)	(1.05)	(1.49)
Reserves-to-loans	0.606	-0.984*	-0.0578	0.456	0.738	0.456	0.463***	-1.124*	0.727	0.308
	(0.95)	(-1.75)	(-0.81)	(0.73)	(0.89)	(0.73)	(2.94)	(-1.97)	(0.91)	(0.51)
Market-to-book	0.0306	0.144***	-0.0035	0.0270	0.153***	0.0222	0.0156	0.137***	0.137**	0.0180
	(0.98)	(2.78)	(-0.61)	(0.86)	(2.66)	(0.72)	(1.55)	(2.70)	(2.44)	(0.58)
Leverage	0.298	-0.0655	0.121*	0.512	-0.218	0.391	0.120	0.0414	0.136	0.520
	(0.37)	(-0.13)	(1.76)	(0.61)	(-0.29)	(0.48)	(0.73)	(0.08)	(0.17)	(0.62)
ROA	0.823	-2.660	0.0418	1.173	-0.663	1.099	0.929	-2.527	-0.315	1.063
	(0.30)	(-0.64)	(0.12)	(0.48)	(-0.16)	(0.41)	(1.21)	(-0.60)	(-0.08)	(0.46)
Long-term funding	0.138	0.189	-0.0174	0.130	0.0069	0.130	0.0884*	0.145	0.000	0.116
	(1.01)	(1.03)	(-0.44)	(1.00)	(0.03)	(0.94)	(1.80)	(0.80)	(0.00)	(0.91)
Non-interest income	0.0042	-0.0339*	0.0022	0.0097	-0.0375*	0.0046	-0.0019	-0.0334*	-0.0319	0.0136**
	(0.45)	(-1.75)	(1.53)	(1.44)	(-1.80)	(0.52)	(-0.55)	(-1.79)	(-1.62)	(2.14)
Loans-to-deposits	-0.0887***	-0.0364	-0.0066	-0.116***	-0.0668	-0.0870**	-0.0337**	-0.0414	-0.0830	-0.113***
	(-2.68)	(-0.64)	(-1.26)	(-3.54)	(-0.75)	(-2.52)	(-2.30)	(-0.73)	(-0.96)	(-3.35)
Tier1 ratio	0.0279	0.472	0.108**	0.0925	0.0447	0.0579	0.0223	0.599*	0.272	0.0918
	(0.11)	(1.52)	(2.25)	(0.31)	(0.12)	(0.21)	(0.28)	(1.92)	(0.72)	(0.31)
Z-score	0.0017	0.0018	-0.0003	0.0012	0.0034*	0.0015	-0.0002	0.0017	0.0030*	0.0010
	(1.66)	(1.33)	(-1.25)	(1.16)	(1.96)	(1.33)	(-0.62)	(1.32)	(1.80)	(0.88)
Interconnectedness	-0.0255	-0.0420	0.0102	-0.0391	-0.110	-0.0262	-0.0095	-0.0340	-0.114	-0.0339
	(-0.46)	(-0.39)	(0.86)	(-0.74)	(-1.11)	(-0.46)	(-0.66)	(-0.31)	(-1.22)	(-0.65)
Unemployment	1.268**	1.070	0.223**	1.263**	-0.320	1.239**	0.0906	1.141	-0.280	1.241**
	(2.40)	(1.37)	(2.15)	(2.38)	(-0.37)	(2.28)	(0.71)	(1.50)	(-0.32)	(2.24)

Table 2.1 (cont.)

Inflation	1.094 (0.98)	0.275 (0.16)	0.277 (1.07)	1.372 (1.19)	-2.469 (-1.38)	0.782 (0.69)	0.154 (0.52)	0.432 (0.26)	-2.347 (-1.26)	1.102 (0.98)
GDP growth	0.114 (0.17)	-1.420 (-1.34)	-0.0489 (-0.37)	0.292 (0.46)	-1.732 (-1.51)	0.153 (0.23)	0.489*** (2.69)	-1.153 (-1.09)	-1.727 (-1.50)	0.382 (0.60)
HHI	-3.420*** (-4.39)	-0.909 (-0.75)	-0.220** (-2.28)	-3.345*** (-4.50)	-1.605 (-1.46)	-3.437*** (-4.51)	-0.566*** (-2.77)	-0.887 (-0.76)	-1.669 (-1.42)	-3.383*** (-4.58)
Stock market return	0.149*** (2.65)	-0.0577 (-0.73)	0.0002 (0.02)	0.154*** (2.91)	-0.0695 (-0.72)	0.156*** (2.82)	0.0171 (0.89)	-0.0590 (-0.72)	-0.0384 (-0.41)	0.160*** (3.14)
Stock market volatility	-0.0685 (-0.35)	-0.388 (-1.17)	0.0473 (1.20)	-0.174 (-0.90)	0.0190 (0.07)	-0.0043 (-0.02)	0.0297 (0.55)	-0.404 (-1.25)	0.117 (0.43)	-0.106 (-0.54)
Long-term interest	-0.00761 (-1.02)	-0.0259** (-2.13)	-0.0002 (-0.17)	-0.0042 (-0.57)	-0.0161 (-1.55)	-0.0070 (-0.91)	0.0042** (2.48)	-0.0250** (-2.14)	-0.0185* (-1.75)	-0.0041 (-0.54)
Const.	-1.853 (-1.52)	1.592 (0.83)	0.719*** (4.26)	-0.970 (-0.76)	-1.829 (-0.79)	-1.848 (-1.50)	0.192 (0.61)	1.257 (0.66)	-1.886 (-0.82)	-0.961 (-0.76)
Bank fixed effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Time fixed effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
N	1850	1793	1851	1850	1792	1850	1850	1793	1792	1850
R^2	0.075	0.084	0.085	0.079	0.068	0.070	0.086	0.082	0.071	0.077
Adj. R^2	0.051	0.060	0.061	0.055	0.044	0.046	0.063	0.058	0.046	0.054

This table contains regression results for rank correlations between exposure-SRMs. Dependent variable is the respective rank correlation mentioned at the top of each column.A panel approach with fixed effects is applied; t-statistics are given in parentheses. The explanatory variables enter the equation with one lag. Regression coefficients are marked with stars if significant as follows: ***= 1% confidence level, **= 5% confidence level, *= 10% confidence level.

The results from the less volatile subsample analysis reveal a considerably different outcome. Here, several different variables are found to be associated with rank correlations and in particular bank-individual explanatory variables are strongly associated with basically the full set of rank correlations. The coefficient of asset size, for example, is more often significant and positive (four out of ten correlations) compared to three ones in the baseline specification. Similarly, the coefficient of the market-to-book ratio is positive and significant (five compared to four in the baseline), the coefficient of leverage is positive and significant (five compared to one in the baseline), and, finally, the coefficient of z-score is positive and significant (five compared to two in the baseline). In addition, the reserves-to-loans ratio is not only more often found to be significant, the sign of the coefficients is also clearly positive for the less volatile sub-sample (five significant coefficients and zero negative ones compared to two negative and one positive coefficient in the baseline). On the contrary, the loans-to-deposits ratio is only in three cases negative and significant compared to five in the baseline, even though the size of the coefficients is slightly larger in case they are significant. Finally, the Tier 1 ratio is strongly significant and positive in the low volatility sub-sample for five correlations, while it was only positive and significant in two cases in the baseline. This is particularly interesting as the coefficients primarily turn out to be negative and significant in the high-volatility sub-sample for four correlations (compared to two positive ones). Hence, an increase in the regulatory capital base during low volatility times leads to an increase in rank correlations, while an increase during high-volatility times induces a decrease. In terms of economic significance, the increase could amount, on the one hand, to 0.03 for the correlation between MES and ΔCoVaR-EQ during a low-volatility regime and, on the other hand, up to -0.02 during a high-volatility regime.

Macroeconomic variables do not seem to be too important during the less volatile periods, even though unemployment, HHI and stock market returns are still found to be significant in several specifications. However, the number of significant coefficients declines and, in the case of HHI, the size of the coefficient decreases as well. Taken together, this can be interpreted as an indication that rank correlations are more strongly influenced by bank-individual variables in periods of low financial market volatility. In general, the larger a bank is and the more stable it appears (based on the reserves-to-loans ratio and the Tier 1 ratio), the stronger SRM ranks tend to be correlated. On the contrary, more indebted banks seem to exhibit higher rank correlation during these market phases, too. The economic significance of this effect amounts to up to 0.05 following a one-standard deviation increase in leverage for the correlation between MES and

ΔCoVaR-EQ.

Table 2.2: Summary statistics - rank correlations in different volatile market phases

		Subsample VIX –lower 50% quantile ('less volatile phase')	Subsample VIX – upper 50% quantile (more volatile phase')
Corr MES-SRISK	Mean	0.3753**	0.4081
	Median	0.4448	0.4765
Corr MES-MES-EQ	Mean	0.6869	0.6808
	Median	0.7565	0.7605
Corr MES-ExpΔCoVaR	Mean	0.9464***	0.9556
	Median	0.9667	0.9707
Corr MES-ExpΔCoVaR-EQ	Mean	0.6832	0.6869
	Median	0.7518	0.7693
Corr SRISK-MES-EQ	Mean	0.2770*	0.3058
	Median	0.3455	0.3432
Corr SRISK-ExpΔCoVaR	Mean	0.3706**	0.4024
	Median	0.4381	0.4663
Corr SRISK-ExpΔCoVaR-EQ	Mean	0.2735**	0.3111
	Median	0.3322	0.3681
Corr MES-EQ-ExpΔCoVaR	Mean	0.6778	0.6735
	Median	0.7468	0.7573
Corr MES-EQ-ExpΔCoVaR-EQ	Mean	0.9022	0.9065
	Median	0.9292	0.9345
Corr ExpΔCoVaR –ExpΔCoVaR-E(	Mean	0.6349	0.6837
	Median	0.7481	0.7711

This table displays the sample mean and median of rank correlations for the sample in less volatile market phases (VIX - lower 50% quantile) and in more volatile market phases (VIX - upper 50% quantile). Quantiles are computed based on the values of the CBOE Volatility Index (VIX). Significant differences between the two subsamples based on a t-test are marked as follows: ***= 1% confidence level, **= 5% confidence level, *= 10% confidence level.

Those results reveal several interesting insights. First, it is shown that the consistency of systemic risk ranks behaves quite differently with respect to different market conditions. Rank correlations are slightly larger during more volatile market phases, but the association to bank-individual variables is considerably weaker. The effect of macroeconomic variables tends to be ambiguous with a subset of rank correlations exhibiting a stronger association with rank correlations during more volatile times, while their influence is limited during low-volatility periods. On the contrary, during low-volatility markets, rank correlations are somewhat smaller but they are to an increasing extent associated with different bank-individual variables.

This is an important finding with respect to a better understanding of the reliability of systemic risk rankings. Since systemic risk rankings are particularly important when market conditions deteriorate, it is also a rather pessimistic insight, as we can see that the sensitivity of systemic risk ranks does not depend on the given set of variables. The correlations of systemic risk ranks appear to be rather unaffected by standard bank-individual variables during stress periods, so it is difficult to relate them with factors that could contribute to or explain their variation, in particular in times when their accuracy is of utmost importance, even though ranks tend to be more aligned in these market phases.

2.5.3 Panel regressions based on weekly systemic risk measures

Until now, SRMs are on a daily basis and the subsequent correlations are computed based on the ranks of daily SRMs within one quarter. Daily values, however, could contain a considerable amount of noise in the underlying return series, such that we also consider weekly SRMs and, hence, ranks based on weekly data. This exercise could allow increasing the robustness of our results beyond results that could be driven by spurious noise. In this case, the number of ranks within one quarter is quite low, such that correlations are based on all weekly ranks within half a year. In a next step, these bi-annual correlations are regressed on the same set of explanatory variables as before, but on a bi-annual frequency as well. The regression results can be found in Table A.7 (in Appendix A). Of course, the number of observations is reduced by 50%, but we can see that the overall results remain similar, even though the significance in general is somewhat lower than for the daily case. Bank-individual variables are mostly found to be insignificant, with once again the market-to-book ratio (but also the reserves-to-loans ratio and in one case the loans-to-deposits ratio) as an exception. In comparison to the baseline, we find that the association between bank-individual variables and rank correlations is reduced and we note that the size of the coefficients is mostly unchanged.

Macroeconomic variables are often significantly associated with rank correlations, this holds once again in particular for unemployment and the HHI. The coefficient of stock market return is still found to be significant for several rank correlations, but to a weaker extent.

These findings show that the overall picture that we have seen so far is somewhat robust towards the frequency of the data of the underlying SRM computation. It is also an indication that bank-variables are to a stronger degree affected by potential noise in the underlying daily series, while the effect of macroeconomic variables is still similarly strong, or even stronger, as

the size of the respective coefficients for unemployment and HHI tend to be slightly larger than in the previous setting. On the contrary, some variables such as the long-term interest rate are not found to be associated with rank correlations in this set-up anymore, while the association to stock market volatility turns out to be strongly significant now. This can be interpreted, on the one hand, as good news, given that it becomes clearer that some relations can be interpreted more reliably (those of unemployment and HHI). On the other hand, this is rather bad news as it shows that bank-variables indeed do not have a particularly pronounced effect on rank correlations and that not all associations with macroeconomic variables remain stable, hindering a straightforward interpretations of different SRMs.

2.6 Alternative cross-sectional analyses

In addition to the estimation of the panel regression specifications, another type of correlation is computed and further regression analyses are conducted to shed more light on the behavior of rank correlations.

While so far bank-individual rank correlations in a time-series dimension have been considered, it is also interesting to study the rank correlation for a given point of time (i.e., quarter) across all banks (as done by Nucera et al. (2016), Benoit et al. (2019) and Grundke (2019)). Hence, it is no longer studied how the ranking consistency of specific banks is influenced in a time-series dimension, but if there are common patterns in the behavior of cross-sectional rank correlations. For each quarter of the sample period, Spearman's rank correlation between any pair of the considered SRMs during this quarter is computed.[9] For this, first, rank correlations between pairs of SRMs across all banks on one day are computed and, second, the average of the daily rank correlations during one quarter is determined. In addition, the explanatory variables are averaged over all countries for the given (quarterly) period. This analysis is conducted for the whole sample as well as for different groups of countries in a further step, too, in order to

[9] As an alternative to Spearman's rank correlation, Kendall's rank correlation, also known as Kendall's tau coefficient, can be computed, too. The basic idea behind this ranking is not to measure the distance between the ranking outcomes of pairs of ranks but to simply measure whether or not the ranks of a series deviate in the same direction or not. This leads to a more robust procedure towards outliers. Therefore, Kendall's rank correlation coefficient is found to compute correlations that are more centred around zero compared to Spearman's rank correlation coefficient. The subsequent analysis is repeated with Kendall's tau instead of Spearman's rho. Results (not shown) are qualitatively similar.

investigate whether or not the rank consistency of SRMs show differentiated patterns in different countries.

For each pair $d_1, d_2 \in \{1, \dots, D\}$ of SRMs, the regression equation is as follows:

$$Corr\left(R_{d_1}^q, R_{d_2}^q\right) = \beta_0 + \sum_{p=1}^{P} \beta_p^{d_1,d_2} \cdot Macroeconomic\ variables_{p,\emptyset,q-1} + \varepsilon_q^{d_1,d_2}. \quad (2.7)$$

Table A.8 (in Appendix A) shows summary statistics for each considered pair of rank correlation.

Kendall's coefficient of concordance (see Kendall and Gibbons (1990), p.119) is employed as an additional overall estimator of rank consistency. It is computed as:

$$kconc_q = \frac{(5-1) \cdot Avg\left(Corr\left(R_{d_1}^q, R_{d_2}^q\right)\right) + 1}{5} \quad (2.8)$$

with *Avg* indicating the average of the subsequent correlations across all different pairs of SRMs. It measures the degree of agreement of ranks based on different SRMs. Its values range from zero to one with larger values indicating a stronger degree of agreement.

The correlations of ranks based on different exposure SRMs across all banks vary considerably, similar to the time-series correlations computed before. The correlations between MES and ΔCoVaR are once again very high with values around 0.9, but only for correlation between the "pure" version and the correlation between both variants with equity. Multiplying one of the SRMs with market value of a banks' equity decreases not only the correlations tremendously, but the share of significant correlations among all daily correlations also shrinks from 100% to only around 20-30%. Correlations that contain SRISK ranks are low for MES and ΔCoVaR (0.21 and 0.27), but higher for MES-EQ and ΔCoVaR-EQ (both around 0.6) while they are significantly different from zero for the majority of correlations.[10]

Turning to the results of the regression illustrated in Table 2.3, it can be seen that particularly the HHI as a measure for market concentration of a country's banking industry, a country's stock market volatility and a country's long term interest rates are found to be related to the cross-sectionally measured rank correlations of SRMs. However, the coefficients for HHI are

[10] Those results are in line with Benoit et al. (2013) who find in related fashion correlations between ranks of MES and SRISK to be considerably smaller than between MES and (contribution-) ΔCoVaR. However, they do not test for rank correlations stemming from SRMs that are computed additionally based on banks' equity market value.

not consistent; approximately half of them are positive while the others exhibit a negative sign (regardless of being statistically significant or not). It is therefore difficult to assess the explanatory power of market concentration on cross-sectional rank correlations even though previous results from the bank-individual panel regressions are somewhat confirmed, as HHI is still found to be significantly associated with rank correlations. The ambiguous results with respect to the HHI might also be attributed to the different facets of market concentration as put forward in Beck et al. (2018). Nevertheless, it is an indication that the relation between market concentration and systemic risk is not fully captured up to now.

The coefficients of stock market volatility and to a minor degree those ones of the long-term interest rate are mostly positive whenever they exhibit a significant coefficient. If increased sovereign long-term interest rates are interpreted as a proxy for increased uncertainty on the country level (similar to the effect of stock market volatility), this finding indicates that systemic risk measures tend to be more consistent in a macroeconomic environment that is marked by such a higher uncertainty. Nevertheless, this tentative conclusion needs to be taken with care as many coefficient are estimated to be insignificant. This result is partly in line with our previous findings indicating that unemployment as a proxy for economic well-being on the country level has a strong measurable influence on time-series rank correlations. However, unemployment is not found to be significantly associated with cross-sectional rank correlations within this cross-country analysis. The economic significance of stock market volatility is also limited: a one standard deviation (0.07) increase could at most lead to an increase of rank correlations between MES-EQ and SRISK of around 0.08. With a view on the sovereign long-term interest rate, the economic effect is larger as a one standard deviation (1.7) increase could lead to an increase of the rank correlation between MES and ΔCoVaR-EQ by 0.12.

By turning the perspective towards the individual regressions, it can be seen that most of the correlations are in general not well explained by the tested set of explanatory variables. The rank correlation between MES and ΔCoVaR is the only one that exhibits significant coefficients for four different explanatory variables, while the remaining correlations can only be significantly associated with two or less explanatory variables. A similar picture emerges for Kendall's coefficient of concordance. Not a single factor is found to be significant. However, one has to keep in mind that this result might also be driven by the model set-up. Here, we try to fit a model based on a small set of data with a considerable amount of explanatory variables. With macroeconomic variables available only at a quarterly frequency, this is a drawback that we have to take into account. Unreported additional analyses based on a larger set of daily data (in

particular interest rates) show that the association between rank correlations and (daily) macroeconomic variables tends to be stronger in such a setting.

Given that our sample consists of banks from several different countries, it is interesting to see if there are specific patterns with respect to cross-sectional rank correlations for different economies. An analysis on individual country-level, however, is not meaningful since there are some countries with only one or two banks. Therefore, we aggregate countries to six different country groups that are marked by geographic proximity. The six groups are formed by banks from the USA and Canada (Group 1); the UK and Ireland (Group 2); Germany, Switzerland and Austria (Group 3); France, Belgium and the Netherlands (Group 4), Portugal, Spain, Italy and Greece (Group 5); and finally the Nordic countries Norway, Sweden and Denmark (Group 6). This allows a more nuanced analysis that might deliver valuable additional insights. In a straightforward manner, ranks, correlations and aggregated macroeconomic variables are computed for each group of countries separately.

The results of the regressions can be seen in Tables A.9-A.14 (in Appendix A). Notably, there are distinctly different patterns for the various country groups, and it becomes clear that the overall results discussed above do not hold within each of the groups.

To begin with, unemployment appears to be an important driver of cross-sectional rank correlation for several groups such as Group 1, 2, 4, 6 and to a minor degree Group 5 even though is not found to be significant in the overall sample. The reason can be seen when we take the sign of the coefficients into account. While it is positive for Groups 1 and 4, it is rather negative for Group 5 but with mixed signs for Groups 2 and 6 (correlations of Group 3 remain unaffected by unemployment). Hence, it cannot be concluded that unemployment has a distinct overall effect, but it rather seems to depend crucially on the specific countries. A similar picture with a seemingly national pattern emerges for the relation between HHI and rank correlations in the cross-section. Groups 3, 5 and 6 show strongly significant and negative coefficients for HHI while Group 4 exhibits a positive relation between market concentration and rank correlations. In addition, rank correlations in Groups 1 and 2 appear to be unaffected. As could be seen in the overall aggregate analysis, stock market volatility and sovereign long-term interest rates are also found to be significantly associated with some rank correlations. However, within this more granular analysis, it turns out that the (positive) effect of stock market volatility is only attributed to Groups 2, 3 and to a minor degree 6. For Group 1, on the contrary, a changing sign

is detected such that it is not unanimously clear if stock market volatility has a positive or negative relation. Further insights can also be gained when looking at the coefficients for the sovereign long-term interest rate. A significantly negative coefficient is often found for Groups 1, 2, and 4, while Group 3 mostly exhibits a positive sign. Depending on the specific group, further variables are partly found to be significant, such as inflation or GDP growth rate.

It is evident that the influence of specific variables is not very persistent across the different country groups. More interestingly, various variables have a distinctly different relation within different groups, such that their influence at least cannot be interpreted as being universal. Hence, it can only be stated that there is partly a very strong relation for example between unemployment and rank correlations, but that it is unclear how the relation exactly materializes. Furthermore, stock market volatility on average can be significantly associated with cross-sectional rank correlations with a positive sign; however, there are also several negative coefficients. Still, the broader picture reveals that increased stock market volatility tends to relate to higher correlations, on the country-level as well as for the aggregate analysis.

To sum up, the findings from the aggregate analysis are partly confirmed when looking at specific regional country groups, in particular as far as the influence of stock market volatility is concerned. Moreover, the ambiguous results on the country level might be a potential explanation for the ambiguous results on the aggregate level. Given the different effects in different groups, we can also understand better why many coefficients with changing signs are found in the aggregate analysis (such as for HHI), or why for example unemployment appears insignificant even though it is highly significant within several countries. The influence of specific macroeconomic factors seems to be quite different within different regional groups such that it might be important to understand the underlying processes better. In general, it can be stated that unfavorable economic conditions in terms of a higher unemployment rate and higher stock market volatility seem to boost rank correlations for several countries, but the channels might differ from country to country.

Table 2.3: Regression results for cross-sectional rank correlations

	MES-MES-EQ	**MES-SRISK**	**MES-ExpΔCo-VaR**	**MES-ExpΔCo-VaR-EQ**	**MES-EQ-SRISK**	**MES-EQ-ExpΔCo-VaR**	**MES-EQ-ExpΔCo-VaR-EQ**	**SRISK-ExpΔCo-VaR**	**SRISK-ExpΔCo-VaR EQ**	**ExpΔCo-VaR-ExpΔCo-VaR-EQ**	**K_Conc**
Unemployment	-1.164	-0.373	-0.356	-0.888	0.0466	-1.411	0.0339	0.0204	0.0884	-1.267	-0.0679
	(-0.29)	(-0.06)	(-1.05)	(-0.22)	(0.02)	(-0.34)	(1.16)	(0.00)	(0.04)	(-0.30)	(-0.03)
Inflation	-4.676	3.570	0.711**	-5.066	1.931	-4.010	0.0233	3.524	1.614	-4.515	-0.682
	(-1.40)	(0.77)	(2.81)	(-1.50)	(0.94)	(-1.16)	(1.08)	(0.82)	(0.79)	(-1.30)	(-0.40)
GDP growth	0.564	-0.691	-0.0261	0.771	-0.903	0.149	-0.0132	-0.691	-0.720	0.366	-0.0795
	(0.44)	(-0.34)	(-0.24)	(0.58)	(-1.07)	(0.11)	(-1.54)	(-0.36)	(-0.88)	(0.27)	(-0.11)
HHI	19.85*	-22.46	-2.108*	19.25*	-14.25**	1.01**	0.0443	-22.09	-14.46**	20.77*	1.052
	(2.07)	(-1.49)	(-2.07)	(2.00)	(-2.30)	(2.11)	(0.52)	(-1.57)	(-2.44)	(2.08)	(0.22)
Stock market return	-0.0944	-0.179	-0.0224***	-0.100	0.0946	-0.100	0.0000	-0.168	0.0877	-0.103	-0.0425
	(-0.82)	(-1.11)	(-2.92)	(-0.85)	(1.45)	(-0.85)	(0.05)	(-1.10)	(1.38)	(-0.86)	(-0.75)
Stock market volatility	0.0306	-0.594	-0.0058	-0.0080	1.094***	-0.0018	0.0059**	-0.608	1.064***	-0.0540	0.0890
	(0.08)	(-1.28)	(-0.27)	(-0.02)	(4.93)	(-0.00)	(2.81)	(-1.35)	(4.98)	(-0.14)	(0.46)
Long-term interest	0.0675**	-0.0237	-0.0082**	0.0723**	-0.0268	0.0644**	-0.0000	-0.0274	-0.0261	0.0704**	0.0162
	(2.30)	(-0.45)	(-2.29)	(2.47)	(-1.09)	(2.19)	(-0.28)	(-0.58)	(-1.08)	(2.40)	(1.03)
Const.	-1.547*	2.530**	1.210***	-1.515*	1.504**	-1.612*	0.990***	2.500**	1.539***	-1.592*	0.423
	(-1.91)	(2.21)	(14.21)	(-1.89)	(2.75)	(-1.94)	(151.78)	(2.37)	(2.92)	(-1.93)	(1.16)
N	29	29	29	29	29	29	29	29	29	29	29
R^2	0.376	0.355	0.559	0.377	0.854	0.351	0.563	0.363	0.852	0.366	0.187
Adj. R^2	0.168	0.141	0.412	0.170	0.805	0.135	0.417	0.151	0.803	0.155	-0.084

Dependent variable is the (cross-sectional) rank correlation between the various exposure SRMs based on Spearman's correlations coefficient and Kendall's coefficient of concordance, respectively, with t-statistics given in parentheses. An OLS approach with heteroscedasticity-robust standard errors is applied. Regression coefficients are marked with stars if significant as follows: ***= 1% confidence level, **= 5% confidence level, *= 10% confidence level.

2.7 Discussion

The extensive analyses have revealed several important insights with respect to the relation between systemic risk rank correlations and associated bank-specific and macroeconomic factors. Generally, it can be seen that rank correlations differ to a considerable extent depending on which SRM pair is regarded and depending on the given specification, that is whether they are computed within a time-series or a cross-sectional dimension.

With respect to the full sample, it is hard to find specific factors that can be unanimously associated with rank correlations. Bank-individual variables are particularly bad in explaining variations in rank correlations, even though potentially more stable banks (in terms of a higher market-to-book ratio and lower loans-to-deposits ratio) and larger banks (in terms of their asset size) can be suspected to exhibit higher rank correlations. In addition, rather negative macroeconomic conditions (in terms of the unemployment rate) and lower market concentration in a country's financial industry are found to be positively associated with rank correlations. By computing SRMs on a weekly basis instead of a daily basis, we can show that the influence of macroeconomic variables can be suspected to be more robust, as coefficients for bank-individual variables are hardly found to be significant, contrary to those of macroeconomic factors.

When splitting the sample into a financially distressed subsample and a financially tranquil one, it becomes clear that the influence of the tested set of variables is very different. Rank correlations are larger during more volatile market phases, but they are more weakly associated with any bank- or macroeconomic variable. A small sub-set of rank correlations is more strongly associated with some macroeconomic factors, but this finding does not hold in general. On the contrary, during low-volatility markets, rank correlations are somewhat smaller but they are to an increasing extent associated with different bank-individual variables. This is a very important finding, given that financially stressed market phases are the times when the identification and ranking of financial institutions with respect to their systemic risk potential is particularly important. In other words, systemic risk ranks and, hence, the assigned systemic risk potential can be more reliably assessed during more volatile market periods, because different systemic risk measures tend to deliver similar results. On the contrary, these circumstances do not allow for a meaningful analysis of other factors that contribute to or explain variations of rank correlations. Bank-individual variables are only significantly associated with rank correlations when markets are rather calm. Whenever financial markets are marked by higher volatility, the explanatory power of most of these factors diminishes. Our results suggest that no unique single factor can explain a larger fraction of the variation of SRM rank correlations.

SRM ranks of more stable banks tend to be estimated more similarly, but this finding in not robust across all the different specifications. Macroeconomic variables seem to be a more important factor that could be associated with rank correlations. Those results are not fully robust throughout the different specifications; however, the general trend is observable. Still, it remains obvious that larger parts of the variation of rank correlations cannot be explained.

One has to keep in mind that a significant relation between, say, the unemployment rate and the SRM rank correlation does not imply that unemployment induces higher systemic risk, but it is an indication that the systemic risk potential of banks can be more easily and consistently assigned when unemployment is higher.

As a next step, we shifted our focus to the aggregate rank correlation of SRMs across several banks for a given point in time. This analysis is performed both for the overall sample as well as for subsamples of country groups. It can be stated that some general findings hold quite stable across the different specifications while others solely realize for individual country groups. The sovereign long-term interest rate, the HHI and the stock market volatility are most adequately able to explain variations in cross-sectional rank correlations. The results with respect to market concentration, however, are not clear. For some rank correlations, we find a strongly negative effect regarding market concentration, while others turn out to be strongly positively associated. Hence, we can only conclude that there is an important relation towards market concentration without explicitly deciphering the concrete channels. Our results also suggest that macroeconomic factors affect systemic risk ranks quite differently across different regions. Unemployment is strongly positively associated in some countries while it is strongly negatively associated in other country groups. These findings shall lead to an even more careful interpretation of systemic risk ranks in general. This cautious view is further confirmed when comparing the coefficients of macroeconomic variables between the panel regression and the cross-sectional one. In most cases, the association between SRM rank correlations and macroeconomic variables differs with respect to both significance and the sign of the coefficient. While in particular unemployment, HHI and stock market returns are significant in the panel setting, only HHI is also found to be significant for a number of cases in the cross-sectional setting, but as previously discussed, with changing signs. Others, such as the long-term interest rate is mostly negative and significant for some rank correlations in the panel framework while it is rather positive and significant in the framework of cross-sectional rank correlations.

The partly weak consistency between the rankings of different SRMs could pose considerable

problems to researchers working with those measures. To be more precise, it is unclear whether results based on one or two particular SRMs could hold when some other measures deliver quite different outcomes in terms of systemic risk potential. Bank-individual and macroeconomic variables are not able to explain those differences consistently. Thus, it seems to be rather difficult to identify some kind of 'safe harbor' (given by specific constellations of explaining variables), in which the SRM choice is likely not to influence the results of an analysis. Those findings are confirmed with additional, yet unreported, results. It turns out that the results are quite strongly dependent on the specific setting that is analyzed and the respective computational methods involved. Altering the method to compute the different SRMs by considering, for instance, quantile regressions instead of a DCC-GARCH specification to compute ΔCoVaR or by incorporating other financial indices as a proxy for the market instead of the MSCI world, leads to different results (not reported). In addition, analyzing ranks and rank correlations of contribution SRMs or a mixture of exposure and contribution measures reveals once again rather different results compared to ours. It is suggestive that systemic risk is potentially not fully captured by a single SRM. Given that these measures are often used interchangeably there is a considerable need for a careful interpretation and it is strongly advisable to use many different (in terms of construction principle and in terms of estimation method) SRMs for any sound analysis.

2.8 Conclusions

The need to rank financial institutions based on their systemic risk potential became apparent in the aftermath of the financial crisis. Ongoing efforts have led to the development of sophisticated measures for systemic risk that have been shown to deliver valuable insights with respect to a bank's systemic importance. However, when ranking banks based on different systemic risk measures, ranks turn out to differ largely across those measures. In order to shed light on the diverging ranking outcomes, in this study, we look for explanatory factors for the correlations of ranks. In previous studies, a large set of bank-specific and macroeconomic factors have been exploited in order to explain the level of systemic risk measures. We embark on this strategy and investigate similar variables with respect to their relation towards rank correlations. A panel data approach detects only few explicit bank characteristics that are significantly associated with the dynamics of rank correlations. Macroeconomic variables such as the unemployment rate or the Herfindahl-Hirschman-Index (HHI) as a proxy for the market concentration of a country's financial industry are more strongly and more often significantly associated with

rank correlations. This finding is partly confirmed when repeating the analysis for periods of particular financial distress and tranquillity. It turns out that most rank correlations are seemingly unaffected by any bank-individual variables during high-volatility market phases. Neither can macroeconomic variable be associated with most rank correlations with some notable exceptions for rank correlations of SRMs that are computed with a variant of a bank's market value, while the same set of variables has significant explanatory power during periods of low-volatility. This result can help practitioners and academics to understand the rankings of systemic risk measures better.

An aggregate analysis that measures rank correlations across banks for a given point of time strengthens the view that it is very hard to find specific factors that have a robust influence on rank correlations and it can be shown that these findings are strongly dependent on the set of countries under investigation. Extensions and further robustness tests confirm these results and they show that only a small part of the variation in rank correlations can indeed be explained by the regarded variables. Thus, it can be suspected that the underlying process is more complex than assumed so far. More sophisticated analyses need to be conducted to elaborate further on rank correlations of systemic risk measures. Those results will enable e.g. regulators and policy makers to assign a bank its individual risk potential more precisely and could hence reduce the fragility of the global financial system as a whole.

Contributions to the second essay "From cashtag to hashcrash – Predicting financial market volatility with Twitter" by Michael Abendschein and Gibran Watfe

The initial research proposal has been jointly developed by Gibran Watfe and Michael Abendschein. They have also jointly contributed to the implementation of the research design.

Gibran Watfe has to a major extent carried out the statistical analyses with respect to forecasting daily and intradaily volatility and he contributed to drafting the manuscript.

Michael Abendschein has to a major extent been responsible for developing the research design, data gathering, and the analysis and presentation of results in the manuscript.

Michael Abendschein Gibran Watfe

3 From cashtag to hashcrash – Predicting financial market volatility with Twitter (Michael Abendschein and Gibran Watfe)

3.1 Introduction

For many areas of the economy, online data is an ever more important source of information and allows for novel insights into the behavior of economic agents. This holds in particular for financial markets. Nowadays, a large share of financial market transactions is automated. The underlying algorithms take decisions using data inter alia from online sources. The messaging service Twitter constitutes one of the most frequented sources of online information. The significant role of Twitter data for financial markets made headlines, for instance, on 6 December 2016. In a tweet, US president-elect Donald Trump alleged that costs of the new government commissioned Boeing jet were "out of control" and called for a cancellation of the order (Revesz 2016). The remarks triggered a selloff of Boeing stock in the magnitude of 1 billion USD within a few minutes which was reversed on the same trading day. The effect of Twitter activitity hence did not overly affect Boeing stock's daily returns but their intraday and daily volatility. This - admittedly exceptional - episode demonstrated that equity investors do respond to activity on Twitter and that there is a rational to investigate the usefulness of Twitter for the prediction of financial volatility.

A large body of empirical literature looks at the relationship of various forms of media-generated content and financial market developments. Early research analyzed the information value contained in newspapers for financially relevant news, e.g. Tetlock (2007). The advent of the internet was accompanied by research interested in the influence of chat rooms and online fora on financial market movements, e.g. Antweiler and Frank (2004). In recent years, widespread online services such as Google Search and Twitter have become the main focal points for research in this area.

Data originating from Twitter is increasingly used to analyze financial market developments. Research mainly looks at three variables, trading volume, stock returns and stock volatility. The extent to which Twitter activity relates to these measures varies. Correlations have been found to be quite robust for trading volume. The most prominent studies focus on stock returns as there is pecuniary potential if correlation or causation between Twitter activity and stock returns can be established, see Mao et al. (2011) and Mao et al. (2015). However, the link between the two seems to be rather weak. Finally, stock volatility has received increasing attention in recent years. The first studies indicate a positive correlation and some signs of causation from Twitter

activity to stock return volatility, see Sprenger et al. (2013). Behrendt and Schmidt (2018) find that Twitter might have a positive effect on intraday volatility forecasting, but that this effect is marginal and insignificant.

These diverging results highlight the necessity for a comprehensive study regarding the ability of Twitter data to improve volatility forecasting. We use a large sample of stocks consisting of constituents of the S&P100, all so-called globally systemically important financial institutions as well as some of the major stock indices. Furthermore, we investigate daily and intraday volatility to detect if there is a particularly different pattern. This approach is also driven by methodological considerations. Daily volatility models are heavily used and applied in academia and practice such that its specific application here shall be straightforward. Intraday volatility models, in contrast, are rather rarely employed such that there is a greater variety of approaches. This procedure aims at rendering our findings particularly robust. We first model volatility on a daily level and compute one-step ahead forecasts mainly based on the heterogeneous autoregressive (HAR) model, see Corsi (2009). Subsequently, tweets sampled on five minute intraday intervals are used to compute intraday volatility forecasts based on an ARMA model. Thereby, volatiltity is estimated using various different approaches.

The aim of this work is not to successfully forecast the volatility of a random stock or equity index based on Twitter data or to state a specific statistical relationship. The aim is rather to conduct a comprehensive analysis of a meaningfully large fraction of the equity market in order to understand the overall performance of Twitter data for forecasting stock volatility. Therefore various state-of-the art econometric techniques are applied that allow for a statistically sound and robust analysis. It is important to note that all data employed for this paper is publicly available.

Our results indicate that the usefulness of Twitter data to improve forecasting performance is limited. Even though we find sporadic significant improvements, the volatility of different types in the majority of our sample cannot be significantly more precisely estimated when we include Twitter information. This finding holds across most specifications. Hence, our results confirm to a certain degree previous findings. Most directly, we confirm the results from Behrendt and Schmidt (2018) who cannot find a considerable improvement for constituents from the Dow Jones Index on an intraday level. However, Twitter information can be valuable for certain individual stocks.

Forecasting intraday volatility per se is relevant for various types of financial market

participants. First, it can be used by speculators such as hedge funds, as Engle and Sokalska (2012) point out, for the purpose of intraday risk management. Secondly, empirical evidence of order pricing suggests that temporary intraday volatility influences the choice of limit orders and hence the bid-ask spread, see Ellul et al (2007). Twitter data might also be employed to construct early warning indicators on the level of single stocks and for the wider financial system. In particular, financial regulators might be able to extract signals of financial instability from Twitter-enhanced volatility forecasts. Cerchiello et al. (2017), for instance, use Twitter data to enrich a network model for the analysis of systemic risk spillovers. Since banks are particularly important from the viewpoint of financial stability, the ability of Twitter data to be used in this context is highlighted in the following analysis by including all financial institutions that are considered to be systemically important.

In the following, we first summarize the literature that this paper relates to. Chapter 3.3 describes the data used and the various steps of data collection and transformation. Chapter 3.4 illustrates the computation of realized volatility on a daily level and documents results from daily in- and out-of sample forecasting. Chapter 3.5 mirrors this approach for the intraday analysis. Chapter 3.6 summarizes the main findings of the paper and concludes with an outlook.

3.2 Literature

This chapter briefly reviews the related literature along four dimensions: the theoretical underpinnings, the source of information for financial news, the financial data that is analyzed in relation to the news and measures for information content in the source. Another related strand of literature is research on volatility forecasting. As this plays a central role to the formulation of the methodology, related studies are described in Chapter 3.4.

Economic theory is ambivalent as to whether Twitter activity could have an influence on stock returns. The efficient market hypothesis (EMH) developed by Fama (1970) stipulates that all available information is immediately priced into financial markets. Market movements are unpredictable as no single investor has an information advantage. Thus, prices should follow a random walk process. This leads to the assertion that investors cannot systematically outperform average market returns. Crucially, the EMH is based on the assumption of profit-maximizing and rational investors. However, the EMH is challenged by behavioral observations, as brought forward by Kahneman and Tversky (1979) or De Long et al. (1990). Proponents of this strand of theory argue that stock prices can be predictable to a certain extent

thanks to the presence of common biases among investors. It has been shown, for instance, that investors act in herds, systematically overestimate their personal abilities and tend to over- or underestimate risks depending on the prevailing majority view, see e.g. Allen et al. (2006).

More recently, social mood was shown to be one factor systematically influencing investors in their decisions, see Da et al. (2015). Twitter, in particular, is a widely used platform to share information in condensed form, i.e. limited to originally 140 characters per message. It can be used, inter alia, by investors to express opinions about investments and to share information on economic and company-specific news. Hence, Twitter can be seen as a part of the public sphere where information is exchanged.

Before the widespread emergence of online media, newspaper archives constituted the central source of news used by investors and hence for research purposes. In line with behavioral explanations of financial market movements, market moods conveyed through newspaper articles were hypothesized to affect market developments. In his seminal study Tetlock (2007) uses a widely cited column in *The Wall Street Journal* to measure the extent to which the pessimistic view on market developments expressed in the column is able to predict stock returns based on a history of 15 years of daily data and he finds a quite strong link.

As the first online source of financial news, research made use of data from chat rooms and online fora that specifically discussed particular stocks. Wysocki (2000), for instance, investigates the influence of the attention of participants of the chat room *The Motley Fool* towards a specific stock on its trading volume finding significant correlations. Antweiler and Frank (2004) measure the number of messages on *Yahoo! Finance* and *Raging Bull*, two popular discussion fora for stock market investors at the time. They find relations between the discussants' attention towards specific stocks on the one hand and their returns and volatility on the other hand.

As the internet quickly advanced to host online mass media, including social media, the literature turned towards popular services such as *Google Search*, *Yahoo* and *Twitter* for extracting information on financial news. Arguing on theoretical grounds, De Long (1990) predict that noise trading activities lead to increased stock volatility. Using *Google Search* query data, Dimpfl and Jank (2016) confirm this assertion and argue that this type of data represents retail (i.e. non-professional) investors' attention to specific stock indices. A similar study conducted by Bordino et al. (2012) uses *Yahoo* search query data and comes to the conclusion that activity on *Yahoo* influences stock trading volumes. More recent research

increasingly turns towards Twitter as a source of financial news. Similar to earlier approaches, Twitter activity is seen as a proxy for investors' attention and, uniquely, also for investor sentiment (or mood) towards specific stocks or indices (see, for example, Zhang et al. (2011), Mao et al. (2015) and Behrend and Schmidt (2018)).

Focusing on Twitter data specifically, a number of studies have looked at the impact of online activity on stock trading volumes. Benthaus and Beck (2015), for example, find a statistically significant correlation between the number of tweets containing specific keywords and corresponding stock trading volumes.

Building mostly upon behavioral theories a large body of literature focuses on the informational value of Twitter data for predicting stock returns. Results are mixed as some studies find rather weak statistical links that might not allow for reliable forecasts of stock returns. Simple correlations often yield positive associations between measures of Twitter activity and stock returns (e.g. Mao et al. (2011) and Sprenger et al. (2014)). Controlling for other factors of stock return behavior, however, eliminates the significance of the link in some studies, e.g. Sprenger et al. (2014)). Mao et al. (2011) still find significant influences of Twitter activity on stock returns using regression analyses but they do not find a significant out-of-sample forecasting performance of Twitter activity. Ranco et al. (2015) find rather low correlations and insignificant causal relationships between Twitter activity and stock returns. However, they conclude that the relationship might be time-varying and state-dependent as significant results tend to be found in periods of particularly high Twitter activity. Quite recently, Renault (2017) shows that using a sophisticated sentiment analysis can improve return forecasts on an intraday horizon for the S&P500 index.

While the literature using stock returns as a dependent variable is quite developed, research has focused less intensively on the influence of Twitter activity on the volatility of stock returns. However, in this case, first evidence indicates the presence of a stronger statistical link. Thereby the influence of Twitter activity on volatility is both direct and indirect. In the latter case the effects are transmitted through trading volume of particular stocks. The positive empirical relationship between trading volume and volatility in general is well established in the literature, see e.g. Cont (2001). While studies confirm this indirect link via trading volumes, some studies find evidence for a direct link between stock- or index-specific Twitter activity and respective volatilities (see Bordino et al. (2012) and Mao et al. (2011)). Recently, Fan et al. (2019) examine the potential influence of Twitter bots that generate tweets automatically. Their

findings indicate that both stock volatility and the respective trading volumes can be affected by autonomous algorithm-driven tweets, highlighting the need for a better and more general understanding on the influence and extent Tweets can exert on financial markets. On the contrary, the results from Behrendt and Schmidt (2018) reveal rather pessimistic results when it comes to forecast improvement with Twitter.

A third dimension relevant for the analysis in this paper is the measurement of information content in Twitter messages (tweets). The existing literature measures the information content in two different ways: simple counts of tweets and more sophisticated sentiment analysis.

One measure is the simple count of tweets containing a specific predefined keyword per time interval. The choice of keywords is arbitrary but usually informed by common sense. Popular choices are commonly used names of companies or abbreviations thereof. This may introduce a significant amount of noise in the data as mentions of companies on Twitter might often not be motivated by trading considerations. This holds, in particular, for well-known retail brands such as *Apple* and *Microsoft*. To a certain extent so-called cashtags are Twitter-specific keywords that indicate a relation of the tweet to news relevant for financial markets. They are used as a marker indicating the ticker symbol together with the US-Dollar sign $ and reveal a financial background of the tweet. This feature is widely used in financial research using Twitter data (e.g. Sprenger et al. (2013) and Souza and Aste (2016)). Zheludev et al. (2014) use a combination of keywords, i.e. full company names, and cashtags to count the number of company-specific tweets.

Another way to extract information content of tweets is sentiment analysis. As Zhang et al. (2011), Zheludev et al. (2014 and Renault (2017) show, the precision of statistical links is increased when the content of tweets is analyzed beyond simple counts. Still, the results from Behrendt and Schmidt (2018) do not show such a considerable improvement. In practice sentiment analysis is utilized to distinguish between a positive and a negative tone, in some cases including a neutral tone. Approaches range from simple labelling of positive and negative tones based on financial dictionaries (e.g. Mao et al. (2011) and Mao et al. (2015)) to more sophisticated natural language mood detection algorithms (e.g. Zheludev et al. (2014) and Souza and Aste (2016)).[11] For the purpose of this paper, sentiment analysis has several drawbacks. While the more sophisticated algorithms constitute black boxes that do not allow

[11] The tools used in the above mentioned studies are *SentiStrength* and *PsychSignal*.

for precise conclusions, more simple approaches have many potential tweet-mood specifications whose value added might not be stable over longer horizons. There is no consensus on how to measure sentiment correctly. Therefore we stick to the more objectively measurable variable of Tweet counts.

This paper makes use of data on Twitter activity in the form of predefined company- and index-specific keywords and cashtags to identify potential value added for out-of-sample forecasts of stock return volatilities. This is considered to be useful given that information transmitted via social networks spreads at high speeds. For this purpose the analysis makes use of state-of-the-art volatility forecasting methods that are specifically designed for the use of high-frequency data. This is complemented with rigorous sensitivity analysis of the results using, for example, various measures for realized volatility. Importantly, and in contrast to most of the existing literature in this field, a very wide selection of stocks allows for more general and robust conclusions on the potential of Twitter data for volatility forecasting.

3.3 Data

The dataset used for this paper combines data on Twitter activity and stocks. Each observation links one of 155 constituents with tweets, stock prices and trading volume at a one-minute frequency over the sample period.[12] Constituents include five major equity indices as well as 150 companies. The group of companies included in the dataset is made up of 100 S&P100 constituents, 28 Global Systemically Important Banks (G-SIBs; some of them are also part of the S%P100)) and 33 Domestic Systemically Important Banks (D-SIBs). The sample period starts on 7 September 2016 and runs through 30 June 2017.

The sample selection is based on two criteria. First, the aim of the study is to investigate the potential of Twitter data for volatility forecasting on a large scale. That is, the focus of the paper is not to randomly find a stock or an index that proves to be particularly suitable for volatility forecasting with tweets but to investigate the equity market as a whole. Since Twitter activity is most prominent in its home country, the United States, and since the S&P100 index covers a reasonably large part of the US equity market, it is a natural choice. It can be assumed that Twitter users are mostly concerned with popular and larger companies and the companies forming the S&P100 index appear to be a perfectly representative sample that serve our

[12] The list of constituents is shown in Table B.1 in Appendix B.

purpose. The index covers several industries that might respond differently to Twitter activity and it includes both companies with strong brands that might be heavily discussed in public without a link to financial news as well as companies that are rather unknown in public and whose tweets might be attributed more directly to financial news. Finally, the sample is particularly focused on systemically important banks and hence the financial industry. The reasoning is based on the idea, that the potential of volatility forecasts for banks might be especially important in a financial stability framework where sudden volatility spikes could trigger a cascade throughout the financial system. This view is supported by work from Cerichiello et al. (2017) who formalize the influence of Twitter data on systemic risk and financial stability. Therefore, banks that are classified as G-SIBs or D-SIBs are taken into account. Note that this allows in addition to expand the analysis to non-US companies.

Table 3.1: Summary statistics

	Min	**Median**	**Mean**	**Max**
Trading volume	0.00	3.27	11.05	1167.25
Tweets (cashtag)	0.00	18.00	52.68	5318.00
Tweets (keyword)	0.00	76.00	829.30	132124.00

Summary statistics of daily data for the whole sample on the company level between 7 September 2016 and 30 June 2017. That is, for instance, approximately 53 tweets are published on average per day containing a specific cashtag. Volume: in million shares.

Data on Twitter activity was obtained through the Twitter API (application programming interface) in daily data collection exercises. Through its API Twitter provides only a small sample (i.e. less than 1%) of overall tweets. This is mainly due to bandwidth issues since, on average, around 6,000 tweets are posted per second. However, the sample is randomly selected and thus serves the purpose of our analysis. A program was set up collecting all tweets contained in the randomized sample through the Twitter API that contained either one of the specified cashtags or one of the specified keywords.[13] Each tweet that was collected comes with

[13] A cashtag is indicated by a Dollar symbol followed by the respective ticker symbol. It serves as an indicator that the tweet is explicitly meant to refer to the stock of this company. For example, a tweet containing *$MSFT* can hence directly be counted as a tweet that refers to the Microsoft stock.

a time stamp. In a second step, the program counted the number of tweets over one-minute intervals per specified cashtag and keyword. In total, we collected 7,251,564 tweets containing a cashtag, and 62,301,702 tweets containing a keyword, as shown in Table 3.1[14] The number of tweets is substantially smaller for cashtags with 53 tweets per day on average than for company names with a mean of 829 tweets.

Data on stock prices was obtained from Google Finance. A program scraped the data for low, high, open and close stock prices as well as trading volumes on a one-minute basis for a whole trading day. The data on indices was only available on five-minute intervals. The stock data was merged with the tweet data based on the date-time variable and the constituent.

In order to facilitate the use of all variables for analytical purposes, several transformations on the original data were applied. The number of tweets (both cashtag and keyword tweets) were transformed to $ln(1 + number\ of\ tweets)$.

3.4 Daily volatility analysis

3.4.1 Methodology

Volatility is a latent variable that cannot be observed in reality. The literature contains a wide variety of approaches to measure volatility. In this paper, the realized volatility (RV) estimator as proposed by Andersen et al. (2001) and Andersen et al. (2003) is used as this is the most straightforward and most often used volatility proxy within the financial and econometric literature. The simple daily realized volatility estimator is a sum of squared (log-) returns (difference between log-prices p) over a pre-defined interval:

$$RV_t = \sqrt{\sum_{j=1}^{n} (lnp_{t,j} - lnp_{t,j-1})^2} \tag{3.1}$$

Accordingly, the realized volatility at day t is the square root of the sum of squared (log-)returns $r_{t,j}$ sampled at n intraday intervals j of equidistant length.

The realized volatility estimator is based on the assumption that logarithmic asset prices follow a continuous semi-martingale process. It can be shown (see Andersen et al. (2001) and Hautsch (2012), amongst others) that the realized volatility estimator in equation 3.1 can be a consistent estimator of the so-called quadratic variation, the variation of the assumed price process. In

[14] The number of distinct tweets collected is lower as some tweets contained several cashtags or keywords.

other words, assuming that the price process of our asset prices exhibits the given characterisitics, the realized volatility can be computed as the empirical counterpart of the (theoretical) quadratic variation of the price process by summing up all individual (squared) returns of the intraday intervals. (This procedure is not only applied to daily intervals, i.e. to compute daily volatility, but also to smaller, intradaily intervals as can be seen in the next chapter regarding intradaily volatility forecasts). Increasing the intraday sampling frequency n can lead to a convergence of the RV estimator to the latent volatility (or quadratic variation). However, this is only the case for infinitesimal intraday intervals. In practice a trade-off is present between convergence and stability due to market microstructure noise.[15] Previous research has shown that microstructure noise increases linearly with the sampling frequency n and causes the estimator to be driven exclusively by noise at very high frequencies, see Hansen and Lunde (2006). According to the related literature, the optimum for this trade-off seems to lie at a return sampling frequency of around 5 to 20 minutes. This paper uses 5-minute returns for the computation of the RV estimator. In addition, we use the realized variance (RV^2) and the log-transformation $log(RV)$ in order to strengthen the robustness of our results, similar to e.g. Andersen et al. (2007).

Similar to the computation of volatility proxies, there is a plethora of different volatility forecasting models that are more or less able to capture specific characteristics of volatility and therefore they are employed in order to compute volatility forecasts. In this work, we forecast one day ahead volatility based on the heterogeneous autoregressive (HAR) model. This is a natural choice for two reasons. First, it is a simple to estimate model that does not require a disproportionate computational effort. The aim of this work is not to find a forecasting methodology that might produce most accurate forecasts but to detect general results with respect to the informational value of Twitter data based on widely used and widely respected models. Second, the HAR model has already been successfully used to study volatility forecasts enriched with Google data in Dimpfl and Jank (2016) such that we can assume it might also be useful with respect to Twitter data. It was first introduced by Andersen et al. (2007) and Corsi (2009). The latter explains the idea of the HAR model by distinguishing between different types of investors. Accordingly, investors with short, medium and long-term horizons affect short-term, medium-term and long-term volatility, respectively. Therefore, volatility can be

[15] Market microstructure noise is extensively covered in the literature. It originates inter alia from the discreteness of prices and bid-ask bounces.

decomposed into three parts: short-term, medium-term and long-term volatility. On this basis, the authors defined the three parts as horizons of one trading day, one trading week (i.e. 5 working days) and one trading month (i.e. 22 working days). More specifically, a hierarchical structure is assumed such that short-term investors and hence short-term volatility depends on all three types of volatility whereas monthly volatility only depends on its own values. This represents trading behavior in so far as day traders take longer lasting trends into account whereas long-term investors tend to look through short-term price movements.

For the purpose of this paper different specifications are defined: a benchmark specification with the lagged values of RV estimator only,

$$RV_{t+1} = \beta_0 + \beta^d RV_t^d + \beta^w RV_t^w + \beta^m RV_t^m + \epsilon_t, \tag{3.2}$$

a second specification including the number of Tweets per time interval t containing a respective company-specific cashtag as additional regressor,

$$RV_{t+1} = \beta_0 + \beta^d RV_t^d + \beta^w RV_t^w + \beta^m RV_t^m + \gamma TC_t + \epsilon_t, \tag{3.3}$$

and a third specification including the number of Tweets that contain the company name.

$$RV_{t+1} = \beta_0 + \beta^d RV_t^d + \beta^w RV_t^w + \beta^m RV_t^m + \gamma TN_t + \epsilon_t \tag{3.4}$$

The β coefficients represent the marginal effects of daily (subscript d), weekly (subscript w) and monthly (subscript m) realized volatility on the volatility at $t + 1$. While daily volatility is measured based on the specified volatility estimator, the weekly and monthly estimators are measured based on the average over daily realized volatilities during the respective period:

$$RV_t^k = \frac{1}{s} \sum_{\tau=t-s}^{t} RV_\tau^d \tag{3.5}$$

with $\{k, s\} = \{w, 5\}$ for weekly realized volatility and $\{k, s\} = \{m, 22\}$ for monthly realized volatility.[16] ϵ_t is an iid error term and TC_t and TN_t represent the number of cashtags (TC, Twitter Cashtag) and company names (TN, Twitter Names) per time intervall *t*, respectively.

In addition, as a further specification we incorporate a jump component into the HAR model. It has been shown that the precision of the estimation of volatility can be significantly improved when disentangling a continuous from a jump component in the price process and incorporating this finding into the forecasting model, see e.g., Andersen et al. (2002) or Andersen et al. (2007).

[16] Note that this representation is a simplification of the multi-period realized volatility in line with Corsi (2009).

In order to strengthen the robustness of our results, we also investigate whether or not jump robust models can deliver different outcomes. Following Andersen et al. (2007), our HAR regression equation (in the benchmark case) looks as follows:

$$RV_{t+1} = \beta_0 + \beta^d RV_t^d + \beta^w RV_t^w + \beta^m RV_t^m + \beta^U U_t + \epsilon_t \tag{3.6}$$

with U_t measuring the number of significant jumps within day t. Prices changes are defined as jumps following a procedure based on work from Andersen et al. (2007) and Barndorff-Nielsen and Shepard (2004), among others. By disentangling the price process into a continuos and into a non-continous (jump) component, sudden price movements are defined to be significant following a specific test statistics developped by Huang and Tauschen (2005). We refer to the Technical Appendix B for a more detailed methodological discussion.

A stylized fact of volatility is its long-memory property. Autocorrelations decline slowly requiring fractionally integrated models that are suited to capture this behavior. Even though the employed model does not directly model the long-memory effect, Corsi (2009) and Chiriac and Voev (2011) show that the HAR model captures the long-memory characteristics of return volatility impressively well.

Forecast evaluation on a daily level is pursued following some standard procedures, see Patton (2011), i.e. using the mean squared error loss function

$$L(RV_t, E(RV_t|\mathcal{F}_{t-1})) = (RV_t - E(RV_t))^2, \tag{3.7}$$

the quasi-maximum likelihood loss function

$$L(RV_t, E(RV_t|\mathcal{F}_{t-1})) = \frac{RV_t}{E(RV_t|\mathcal{F}_{t-1}) - ln\left(\frac{RV_t}{E(RV_t|\mathcal{F}_{t-1})}\right)} - 1, \tag{3.8}$$

and the R^2 of the Mincer-Zarnowitz regression

$$RV_t = \alpha + \beta E(RV_t|\mathcal{F}_{t-1}) + \epsilon_t. \tag{3.9}$$

The window for in-sample estimation is fixed to roughly 80% of the available data which coincides with all months up to and including May 2017 for the first one-step ahead forecast. Subsequently the window, while staying fixed in absolute size, rolls over the remaining days in the sample generating one-step ahead forecasts at each point. The result is an out-of-sample series of one-day ahead forecasts of realized volatility for June 2017.

3.4.2 Heterogeneous autoregressive model – In-sample analysis

In order to compare the performance of different models, we look at the summary statistics of adjusted-R^2 of several different regressions. To keep results traceable, we illustrate the mean and the distribution of our sample in Table 3.2. It is subdivided into different types of stocks, different volatility measures and regressions with and without an additional jump component. This procedure allows to clearly distinguish between different specifications.[17] The values cover a wide range from around -0.02 as a minimum to around 0.52 at the maximum end of the scale. On average, the mean is found to be between 0.03 and 0.1. This illustrates the finding that the adjusted-R^2 of many specifications is rather low with several outliers. The most important question, naturally, is whether and to which extent the incorporation of Tweets into the volatility estimation increases the precision of the estimation based on the HAR model. We can see on first sight, that specifications with cashtags show a superior performance in terms of the average of adjusted-R^2s. This finding holds for the vast majority of specifications, hence, the precision of the volatility estimation outperforms the specification with company names and the benchmark case without Tweets for any type of stock, any type of realized volatility estimation and irrespective of the use of jumps. However, theses differences are not statistically significant based on a t-test.

We can see that values for the adjusted-R^2 increase by around 3-10%. Interestingly, a mixed picture results with respect to company names and the benchmark case. Since the adjusted R^2 corrects for additional variables, on average there is no gain from using company names. Next, we can distinguish a clear trend between the three types of stock returns that we consider, the constituents of the S&P100, systemically important banks and some major indices. Index volatilities are on average more precisely measurable than the respective individual stocks.

The differences are large with a mean between 10-15% for index volatilities, 3-10% for S&P100 stocks and only around 2-5% for SIFI stocks. Of course, this result could also be driven by the different sample sizes.

[17]Note, that the log modelling of volatility by construction cannot be carried out for jump specifications because the jump component U can equal to zero.

Table 3.2: Summary statistics adjusted R^2 from different HAR model specifications

no jumps										
		Benchmark			**Cashtag**			**Company name**		
SP100		**Min**	**Mean**	**Max**	**Min**	**Mean**	**Max**	**Min**	**Mean**	**Max**
	RV^2	-0.0173	0.0259	0.2977	-0.0233	0.0301	0.3011	-0.0207	0.0255	0.3069
	RV	-0.0182	0.0586	0.3633	-0.0209	0.0607	0.3919	-0.0205	0.0586	0.4020
	logRV	-0.0150	0.1008	0.4946	-0.012	0.1012	0.5121	-0.0197	0.1008	0.5180
SIFI										
	RV^2	-0.0151	0.0261	0.2183	-0.0195	0.0308	0.2144	-0.0197	0.0275	0.2154
	RV	-0.0145	0.0391	0.1996	-0.0167	0.0432	0.2062	-0.0185	0.0410	0.2008
	logRV	-0.0083	0.0478	0.1634	-0.0138	0.0479	0.1601	-0.0137	0.0497	0.1664
Index										
	RV^2	0.0058	0.0930	0.1667	0.0017	0.0959	0.1631	0.0012	0.0853	0.1631
	RV	0.0738	0.1305	0.1965	0.0950	0.1405	0.1948	0.0938	0.1322	0.1932
	logRV	0.1355	0.1560	0.1985	0.1458	0.1677	0.2004	0.1309	0.1608	0.1951
jumps										
		Benchmark			**Cashtag**			**Company name**		
SP100		**Min**	**Mean**	**Max**	**Min**	**Mean**	**Max**	**Min**	**Mean**	**Max**
	RV^2	-0.0299	0.0438	0.4381	-0.0354	0.0482	0.4352	-0.0350	0.0426	0.4354
	RV	-0.0145	0.0722	0.3884	-0.0151	0.0756	0.3956	-0.0188	0.0717	0.4028
	logRV									
SIFI										
	RV^2	-0.0274	0.0386	0.3347	-0.0313	0.0426	0.3316	-0.0290	0.0400	0.3373
	RV	-0.0188	0.0516	0.1913	-0.024	0.0544	0.2023	-0.0220	0.0543	0.1941
	logRV									
Index										
	RV^2	0.0201	0.1044	0.2032	0.0175	0.1073	0.1990	0.0186	0.1011	0.2011
	RV	0.0983	0.1457	0.2173	0.0977	0.1470	0.2161	0.0968	0.1423	0.2137
	logRV									

This table presents summary statistics concerning the adjusted R^2 of several different HAR-models with and without jump component. No jumps refer to the HAR model. S&P100, SIFI and Index refer to the specific subsamples. Note, that results for logRV with jumps cannot be computed by contruction.

While we only consider the five largest stock indices, there are larger number of individual stocks within the other groups that could explain some of the smaller outliers.

Finally, we can look at differences arising from different technical approaches. While it is not the focus of this work to distinguish between better and worse volatility specifications, it is nevertheless interesting whether or not specific anomalies might be detected. It is obvious that there is a clear ranking between the three types of realized volatility modeling. Best results are achieved with log volatilities, followed by the realized volatility and the realized variance. As Andersen et al. (2007) note, log volatilites are closer to being normally distributed and, hence, their statistical properties are more adequate for the employed statistical analyses. This finding holds robustly across the majority of specifications. Additionally, to incorporate jumps into the HAR regressions clearly increases the precision for any of the different regressions. This result could have been expected with respect to similar results in the literature, see e.g. Andersen et al. (2007) and it clearly materializes in our sample as well.

3.4.3 Heterogeneous autoregressive model – Out-of-sample analysis

A first impression of out-of-sample results can be found in Table 3.3. It shows general results across the whole sample based on the different specifications without jumps. We can see that there is no clear trend and that partly Twitter models outperform the benchmark, but sometimes also the other way round.

To recall the interpretation of the three different forecasting errors, it is important to note that negative differences with respect to the MSE and quasi-maximum likelihood and a positive difference with respect to the Mincer-Zarnowitz regression R^2 indicate a better performance in terms of the Twitter specification. Since typically the absolute values of e.g. the MSE are very small (see e.g. Dimpfl and Jank (2016)) we compute the respective relative change in forecasting errors. Given our larger sample size, once again it is most meaningful to illustrate results with the help of summary statistics.

To begin with, it turns out that the overall picture is rather mixed, indicating that partly both Twitter specifications outperform the benchmark but vice versa we can also see an outperformance of the benchmark. In addition, the considerable spread between minimum and maximum values indicate that large over- and underperformances occur. So there are cases with a tremendous increase or decrease in forecasting precision with the help of either cashtags or company names.

Table 3.3: Out-of-sample forecast evaluation HAR-model without jumps

		Mean	Min.	1. Quan.	Median	3. Quan.	Max.
Improvement MSE Cashtag-BM							
	RV^2	0.1822	-0.1093	-0.0074	0.0033	0.0422	1.7863
	RV	0.0474	-0.0727	-0.0070	0.0018	0.0224	0.4838
	logRV	0.0169	-0.0557	-0.0091	0.0003	0.0176	0.1961
Improvement MSE Company name-BM							
	RV^2	0.0258	-0.0895	-0.0070	0.0019	0.02317	0.2956
	RV	0.0048	-0.0532	-0.0096	0.0004	0.0139	0.0863
	logRV	0.0051	-0.0371	-0.0081	-0.0005	0.0099	0.0948
Improvement Qlike Cashtag - BM							
	RV^2	-0.0012	-0.0162	-0.0024	-0.0001	0.0006	0.0098
	RV	-0.0040	-0.0511	-0.0066	-0.0010	0.0031	0.0266
	logRV	-0.0011	-0.0115	-0.0017	-0.0001	0.0005	0.0051
Improvement Qlike Company name - BM							
	RV^2	-0.0009	-0.0085	-0.0010	-0.0004	0.0004	0.0040
	RV	-0.0013	-0.0380	-0.0055	-0.0000	0.0046	0.0234
	logRV	-0.0001	-0.0068	-0.0010	-0.0000	0.0011	0.0052
Improvement MZ Cashtag - BM							
	RV^2	0.0092	-0.0617	0.5088	0.0003	0.0090	0.1674
	RV	0.0022	-0.0484	-0.0053	-0.0000	0.0042	0.0729
	logRV	0.0006	-0.0419	-0.0048	-0.0001	0.0027	0.0596
Improvement MZ Company name - BM							
	RV^2	0.0020	-0.0407	-0.0084	-0.0195	0.0025	0.0793
	RV	0.0022	-0.0255	-0.0026	-0.0000	0.0040	0.0480
	logRV	-0.0013	-0.0296	-0.0039	-0.0000	0.0026	0.0212

This table shows the results of the out-of-sample forecast analysis.The two models with Tweet component, cashtag and company name, are compared with the benchmark HAR model without jump component. The evaluation is based on three different measures, the mean squared error, the quasi likelihhod measure and the Mincer-Zarnowitz regression. Results are displayed in terms of the relative change of the respective criterion from the benchmark compared to the Twitter model for MSE and Qlike. Note, that positive numbers for these two criteria indicate superior performance of the benchmark model without Tweets, while negative values indicate better forecasts of Twitter models. The evaluation of MZ is displayed in form of the differences of the respective R^2. Positive values indicate a better forecast for Twitter models and vice versa. Note, that results are winsorized at the 5% and 95% level such that parts of the summary statistics are not too strongly driven by considerable outliers.

Table 3.4: Out-of-sample forecast evaluation HAR-model with jumps

		Mean	Min.	1. Quan.	Median	3. Quan.	Max.
Improvement MSE Cashtag-BM							
	RV^2	0.1620	-0.1267	-0.0085	0.0028	0.0630	1.585
	RV	0.0316	-0.0898	-0.0127	0.0012	0.0335	0.3576
	logRV						
Improvement MSE Company name-BM							
	RV^2	0.0298	-0.0941	-0.0100	0.0009	0.0239	0.3604
	RV	0.0127	-0.0699	-0.0102	0.0019	0.0200	0.1777
	logRV						
Improvement Qlike Cashtag - BM							
	RV^2	-0.0020	-0.0175	-0.003	-0.0002	0.0007	-0.0515
	RV	-0.0038	-0.0515	-0.0111	-0.0019	0.0050	0.0332
	logRV						
Improvement Qlike Company name - BM							
	RV^2	-0.0014	-0.0126	-0.0017	-0.0001	0.0003	0.0038
	RV	0.0002	-0.0346	-0.0062	0.0000	0.0085	0.0326
	logRV						
Improvement MZ Cashtag - BM							
	RV^2	-0.0013	-0.0752	-0.0086	-0.0003	0.0034	0.0820
	RV	-0.0018	-0.0572	-0.0089	-0.0007	0.0057	0.0619
	logRV						
Improvement MZ Company name - BM							
	RV^2	0.0004	-0.0400	-0.0029	-0.0000	0.0021	0.0443
	RV	0.0009	-0.0390	-0.0067	-0.0000	0.0088	0.0454
	logRV						

This table shows the results of the out-of-sample forecast analysis.The two models with Tweet component, cashtag and company name, are compared with the benchmark HAR model with jump component. The evaluation is based on three different measures, the mean squared error, the quasi likelihhod measure and the Mincer-Zarnowitz regression. Results are displayed in terms of the relative change of the respective criterion from the benchmark compared to the Twitter model for MSE and Qlike. Note, that positive numbers for these two criteria indicate superior performance of the benchmark model without Tweets, while negative values indicate better forecasts of Twitter models. The evaluation of MZ is displayed in form of the differences of the respective R^2. Positive values indicate a better forecast for Twitter models and vice versa. Note, that results are winsorized at the 5% and 95% level such that parts of the summary statistics are not too strongly driven by considerable outliers.

In addition, we can also highlight the necessity of employing different methodologies.

While the MSE criterion suggests that the benchmark is on average (with respect to both mean and median) better able to capture future volatility, the quasi-maximum likelihood criterion on the contrary shows better results for the Twitter specifications (once again for both the mean and the median) and, finally, the Mincer-Zarnowitz criterion shows different results depending on the specific volatility modelling. However, it is also evident that the differences are very small. Based on the MSE criterion, we find that the median value is 0.33% larger for the cashtag specification than the benchmark while the quasi-maximum likelihood criterion suggest an overperformance of the cashtag model of around 0.01% compared to the benchmark. These stylized results hold across the various specifications.

Hence, based on these general results, we cannot state whether or not the incorporation of Twitter information, either cashtags or company names, might considerably increase the forecasting performance of the employed HAR model. On average, performance improvements and declines cancel each other out, a finding that holds robustly irrespective of the specific volatility modelling.

These mixed results are further confirmed when we look at the jump-robust models as shown in Table 3.4. Similar to before, the MSE criterion indicates an outperformance of the benchmark, the quasi-maximum likelihood criterion the other way round and the outcome of on the Mincer-Zarnowitz criterion depends on the specific volatility modelling. On average, the differences between the benchmark and the Twitter specifications shrinks even more compared to the baseline specifications without jumps.

In a next step, we focus on the out-of-sample performance with respect to our three different sub-samples. The results are displayed in Table 3.5, Table 3.6 and Table 3.7. The largest fraction of our sample consists of the constituents of the S&P100 so it is not surprising to find similar results for this sub-sample compared to the overall results above. Once again, the benchmark HAR model seems to outperform both Twitter specifications with respect to the MSE criterion for most of the chosen volatility specifications. On the contrary, the quasi-maximum likelihood criterion tends to show an outperformance for cashtag models and an underperformance for company name models. Similar to before, the model variants with jumps confirm those results, however, now the improvement or deterioration appears to be slightly stronger than in the case without jumps. The overall conclusion that Twitter enhanced models are not consistently able to enhance the precision of daily volatility HAR models holds for this

sub-sample, too.

The SIFI sub-sample analysis (including 61 banks) reveals partly similar results. Still, the overall picture reveals mixed results, with only partial forecast improvements.While the benchmark model on average outperforms cashtag and company name models with respect to the MSE criterion, cashtag models slightly outperform the benchmark with respect to the median MSE values. This finding, however, is not robust when we take the median values from jump specifications into account as well. Similar to before, cashtag models also outperform the benchmark with respect to the quasi-maximum likelihood criterion, both for both jump-robust and non-jump-robust models. Both mean and median still remain very small. Looking at the third criterion also reveals similar results so that, generally speaking, no distinct differences between the S&P100 and the SIFI sub-sample can be stated.

The third sub-sample highlights mixed results once again. Here, company name models on average seem to outperform the benchmark with respect to the MSE, depending on the specific volatility modelling. As opposed to the SIFI case, cashtags do not show a superior forecasting performance here. The quasi-maximum likelihood criterion still indicates an outperformance of twitter models irrespective of the chosen methodology while the Mincer-Zarnowitz criterion indicates an underperformance of Twitter enhanced models, with the exception of company name models including a jump component.

Not only the (relative) size is important in order to analyze forecast performance of different methods, but, of course, also the significance of the respective forecast error comparison. The p-values of the Clark-West-Test are illustrated in Table B.2 in Appendix B.[18]

The results analyzed so far have already shown that the difference with respect to forecasting improvements are on average not very large and are often found to be very close to zero. Indeed, the results of the Clark-West test for significant differences with respect to the MSE criterion reveals, that the vast majority of differences are found to be insignificant at the 5% level. While summary statistics show that all kinds of feasible p-values are found, the clear majority lies

[18] Typically, forecast error comparison techniques test if the forecast error of one model is significantly different from another model. (This is a sort of t-test adapted to the specific statistical properties of the residuals from forecast regressions). The first, and subsequently widely used, test is the Diebold-Mariano test (Diebold and Mariano (1995)). Similarly to Dimpfl and Jank (2016), we use the Clark-West test as an “advanced” version of the Diebold-Mariano test, as shortcomings of the Diebold-Mariano test are meanwhile regularly acknowledged (see Clark and McCracken (2001) and Diebold (2015)).

outside any significance bands. This finding holds in particular for the overall sample, the S&P100 and SIFI sub-samples while the index sub-sample that any of the forecasting differences are found to be significant at all.

3.5 Intraday volatility analysis

The next question we approach is whether or not the accuracy of intraday volatility predictions can be improved by using high-frequency intraday data from Twitter. The speed with which information is spread to investors in times of well-developed online networks raises the question whether a relationship between online activity and financial market volatility is hidden below the surface of the daily frequency. Attention of online users varies day-by-day but arguably also hour-by-hour and minute-by-minute. Similarly, attention of investors to particular stocks, expressed among others by volatility, varies on higher frequencies. Therefore, it might be insufficient to look at volatility on a daily level, but rather depict the potential of Twitter induced information regarding volatility forecasts on an intraday level.

Similar to daily realized volatility, a proxy for intradaily volatility can be computed by summing up squared log-returns of intraday intervals and eventually taking the square-root (see Hautsch (2012), p.197). A 60-minutes volatility estimate, for example, can be computed as the squared sum of six 10-minutes (log-) returns or the squared sum of sixty 1-minute (log-) returns. The smaller the intervals that are summed up, the more precise the volatility estimate will be. For our purpose, we alter the procedure only slightly to mirror and accommodate to the procedures pursued in our reference papers (Bollerslev et al. (2000), Behrendt and Schmidt (2018)): We use intraday volatility as the sum of absolute intrady (log-) returns. The volatility at day *t* and intraday period *j* is defined as (similar to Hautsch (2012))

$$RV_{t,j}^{\mathrm{H}} = \sum_{\eta=1}^{\mathrm{H}} \left| lnp_{t,j,\eta\Delta} - lnp_{t,j,(\eta-1)\Delta} \right|,$$

where *H* equals the number of equidistant intervalls with length $\Delta = \mathrm{H}^{-1}$ within intraday interval *j* and day *t*. In our work, we use absolute five-minute returns as a proxy for 5-minute intraday volatility. A stylized fact of intraday volatility is a distinct seasonality pattern in form of a u-shape across the trading day, that is volatility is higher in the morning, followed by a considerable decline around noon before it starts to rise again towards the end of the day. This finding is regularly confirmed in the literature, see for instance Andersen et al. (2013).

Table 3.5: Out-of-sample forecast evaluation HAR-model – SP100 –

SP100	No Jumps				Jumps			
	Mean	Min.	Median	Max.	Mean	Min.	Median	Max.
Improvement MSE Cashtag - BM								
RV^2	0.2568	-0.1093	0.0089	1.7863	0.2293	-0.1268	0.0050	1.5851
RV	0.0673	-0.0727	0.0052	0.4838	0.0455	-0.0898	0.0016	0.3576
logRV	0.0204	-0.0557	0.0035	0.1961				
Improvement MSE Company name - BM								
RV^2	0.0310	-0.0895	0.0026	0.2956	0.0353	-0.0941	0.0028	0.3604
RV	0.0049	-0.0532	0.0008	0.0863	0.0134	-0.0699	0.0015	0.1777
logRV	0.0030	-0.0371	-0.0003	0.0948				
Improvement Qlike Cashtag - BM								
RV^2	-0.0018	-0.0162	-0.0002	0.0098	-0.0028	-0.0175	-0.0005	0.0044
RV	-0.0049	-0.0511	-0.0010	0.0266	-0.0035	-0.0516	-0.0016	0.0332
logRV	-0.0011	-0.0115	-0.0001	0.0051				
Improvement Qlike Company name - BM								
RV^2	-0.0010	-0.0085	-0.0001	0.0040	-0.0018	-0.0126	-0.0002	0.0038
RV	-0.0009	-0.0380	0.0001	0.0234	0.0013	-0.0346	0.0016	0.0326
logRV	0.0001	-0.0068	0.0002	0.0052				
Improvement MZ Cashtag - BM								
RV^2	0.0126	-0.0617	0.0003	0.1673	0.0008	-0.0752	-0.0001	0.0820
RV	0.0046	-0.0484	-0.0000	0.0729	0.0008	-0.0572	0.0001	0.0619
logRV	0.0028	-0.0419	-0.0000	0.0596				
Improvement MZ Company name - BM								
RV^2	0.0027	-0.0407	0.0001	0.0793	0.0023	-0.0400	0.0000	0.0443
RV	0.0029	-0.0255	-0.0000	0.0480	-0.0390	-0.8953	0.0001	0.0454
logRV	-0.0002	-0.0296	-0.0000	0.0212				

This table shows the results of the out-of-sample forecast analysis for the subsample of S&P100 constituents.The two models with Tweet component, cashtag and company name, are compared with the benchmark HAR model with and without jump component. The evaluation is based on three different measures, the mean squared error, the quasi likelihhod measure and the Mincer-Zarnowitz regression. Results are displayed in terms of the change of in the respective criterion from the benchmark compared to the Twitter model for MSE and Qlike. Note, that positive number for these two criteria indicate superior performance of the benchmark model without Tweets, while negative values indicate better forecasts of Twitter models. The evaluation of MZ is displayed in form of the differences of the respective R^2. Positive values indicate a better forecast for Twitter models and vice versa. Note, that results are winsorized at the 5% and 95% level such that parts of the summary statistics are not too strongly driven by considerable.

Table 3.6: Out-of-sample forecast evaluation HAR-model – SIFI –

SIFI	No Jumps				Jumps			
	Mean	Min.	Median	Max.	Mean	Min.	Median	Max.
Improvement MSE Cashtag - BM								
RV^2	0.0496	-0.1093	-0.0003	1.6184	0.0435	-0.1268	0.0008	1.5850
RV	0.0101	-0.0727	-0.0009	0.4838	0.0072	-0.0898	0.0005	0.2244
logRV	0.0090	-0.0557	-0.0006	0.1961				
Improvement MSE Company name - BM								
RV^2	0.0174	-0.0895	0.0008	0.2955	0.0208	-0.0941	-0.0001	0.3604
RV	0.0054	-0.0531	-0.0000	0.0863	0.0134	-0.0699	0.0023	0.1778
logRV	0.0094	-0.0371	-0.0007	0.0948				
Improvement Qlike Cashtag - BM								
RV^2	-0.0000	-0.0162	0.0000	0.0098				
RV	-0.0017	-0.0511	-0.0001	0.0266	-0.0008	-0.0175	0.0000	0.0044
logRV	-0.0008	-0.0115	-0.0001	0.0050	-0.0041	-0.0516	-0.0014	0.03324
Improvement Qlike Company name - BM								
RV^2	-0.0004	-0.0085	-0.0000	0.0040				
RV	-0.0016	-0.0380	-0.0003	0.0234	-0.0009	-0.0123	-0.0000	0.0038
logRV	-0.0004	-0.0068	-0.0003	0.0051	-0.0016	-0.0346	-0.0012	0.0326
Improvement MZ Cashtag - BM								
RV^2	0.0049	-0.0434	0.0005	0.1674	-0.0039	-0.0752	-0.0004	0.0820
RV	-0.0484	-0.9988	-0.0001	0.0729	-0.0062	-0.0572	-0.0030	0.0619
logRV	-0.0033	-0.0419	-0.0006	0.0596				
Improvement MZ Company name - BM								
RV^2	-0.0007	-0.0407	-0.0005	0.0793	-0.0031	-0.0400	-0.0001	0.0443
RV	0.0015	-0.0255	0.0002	0.0480	0.0006	-0.0390	-0.0005	0.0454
logRV	-0.0030	-0.0296	-0.0001	0.0212				

This table shows the results of the out-of-sample forecast analysis for the subsample of SIFIs.The two models with Tweet component, cashtag and company name, are compared with the benchmark HAR model with and without jump component. The evaluation is based on three different measures, the mean squared error, the quasi likelihhod measure and the Mincer-Zarnowitz regression. Results are displayed in terms of the change of in the respective criterion from the benchmark compared to the Twitter model for MSE and Qlike. Note, that positive number for these two criteria indicate superior performance of the benchmark model without Tweets, while negative values indicate better forecasts of Twitter models. The evaluation of MZ is displayed in form of the differences of the respective R^2. Positive values indicate a better forecast for Twitter models and vice versa. Note, that results are winsorized at the 5% and 95% level such that parts of the summary statistics are not too strongly driven by considerable.

Table 3.7: Out-of-sample forecast evaluation HAR-model – Index –

Index	No Jumps				Jumps			
	Mean	Min.	Median	Max.	Mean	Min.	Median	Max.
Improvement MSE Cashtag - BM								
RV^2	0.1281	-0.0232	0.0447	0.4463	0.1016	0.0210	0.0505	0.2845
RV	0.0556	-0.0583	0.0234	0.2336	0.0200	-0.0561	0.0044	0.1271
logRV	0.0343	-0.0557	0.0277	0.1376				
Improvement MSE Company name - BM								
RV^2	0.0104	-0.0183	0.0136	0.0328	0.0152	-0.0223	0.0012	0.0807
RV	-0.0031	-0.0532	0.0059	0.0289	-0.0125	-0.0483	-0.0051	0.0086
logRV	-0.0025	-0.0371	-0.0010	0.0293				
Improvement Qlike Cashtag - BM								
RV^2	-0.0042	-0.0159	-0.0008	0.0003	-0.0031	-0.0101	-0.0011	-0.0001
RV	-0.0143	-0.0408	-0.0215	0.0265	-0.0067	-0.0316	-0.0142	0.0332
logRV	-0.0038	-0.0088	-0.0057	0.0051				
Improvement Qlike Company name - BM								
RV^2	-0.0004	-0.0010	-0.0003	0.0002	-0.0002	-0.0020	-0.0000	0.0013
RV	-0.0082	-0.0380	0.0007	0.0040	-0.0046	-0.0150	-0.0026	0.0019
logRV	-0.0016	-0.0068	-0.0001	0.0005				
Improvement MZ Cashtag - BM								
RV^2	-0.0167	-0.0617	-0.0160	0.0267	-0.0188	-0.0243	-0.0204	-0.0103
RV	-0.0083	-0.0484	-0.0144	0.0441	-0.0071	-0.0509	-0.0017	0.0259
logRV	-0.0028	-0.0419	-0.0144	0.0596				
Improvement MZ Company name - BM								
RV^2	-0.0082	-0.0384	-0.0074	0.0202	0.0024	-0.0074	-0.0010	0.0188
RV	-0.0039	-0.0217	-0.0055	0.0171	0.0039	-0.0042	-0.0008	0.0212
logRV	-0.0033	-0.02672	0.0009	0.0117				

This table shows the results of the out-of-sample forecast analysis for the subsample of indices.The two models with Tweet component, cashtag and company name, are compared with the benchmark HAR model with and without jump component. The evaluation is based on three different measures, the mean squared error, the quasi likelihhod measure and the Mincer-Zarnowitz regression. Results are displayed in terms of the change of in the respective criterion from the benchmark compared to the Twitter model for MSE and Qlike. Note, that positive number for these two criteria indicate superior performance of the benchmark model without Tweets, while negative values indicate better forecasts of Twitter models. The evaluation of MZ is displayed in form of the differences of the respective R^2. Positive values indicate a better forecast for Twitter models and vice versa. Note, that results are winsorized at the 5% and 95% level such that parts of the summary statistics are not too strongly driven by considerable.

In order to model and forecast intraday volatility, this seasonal pattern has to be taken into account. A common way to deal with recurring seasonality requires the use of a flexible fourier form that disentangles seasonal and stochastic patterns across the trading day. This method is originally developed in Gallant (1981) and adapted to the specific needs of intraday volatility by Andersen and Bollerslev (1998). This procedure results in the estimation of the seasonal component $\hat{s}_{t,j}$ that is used to standardize absolute intraday returns via

$$RV^*_{t,j} = \frac{RV_{t,j}}{\hat{s}_{t,j}}. \tag{3.10}$$

We refer to the Technical Appendix B for a detailed description of the estimation process of the flexible fourier transformation.The de-trended intraday volatility proxy $r^*_{t,j}$ can now be modeled and estimated by an ARMA model in log notation, as put forward by Hautsch (2012).[19] Similar to before, we stick to an ARMA(1,1) model. Given that we have more than a hundred different time-series it will be impossible to find a single best specification for each of them. In order to derive general results, we advocate the use of a straightforward model version. The benchmark model reads as (see Hautsch (2012), p.206)

$$lnRV^*_{t,j} = \alpha + \beta_1 lnRV^*_{t,j-1} + \beta_2 \epsilon_{t,j-1} + \epsilon_{t,j} \tag{3.11}$$

with α being a constant and $\epsilon_{t,j}$ a white noise error term at day t and intraday interval j. Similarly to the daily set-up, we include the number of Tweets based on cashtags and company names as additional regressors, respectively. They are lagged by one intraday period j in order to allow for a potential reaction of markets:

$$lnRV^*_{t,j} = \alpha + \beta_1 lnRV^*_{t,j-1} + \beta_2 \epsilon_{t,j-1} + \gamma TC_{t,j-1} + \epsilon_{t,j} \tag{3.12}$$

and

$$lnRV^*_{t,j} = \alpha + \beta_1 lnRV^*_{t,j-1} + \beta_2 \epsilon_{t,j-1} + \gamma TN_{t,j-1} + \epsilon_{t,j}. \tag{3.13}$$

Forecast evaluation on an intraday level is pursued by means of the the mean squared error loss function

[19] Similar to using the the square-root of squared intraday returns (summed over all intraday intervals) as a proxy for daily volatility, the variation in an intraday interval itself could be estimated as the square-root of the squared individual returns of the given intraday interval. However, in order to stick to the widespread convention in the literature, we use the representation as absolute intraday returns, instead.

$$L(RV^*{}_{t,j}, E(RV^*{}_{t,j}|\mathcal{F}_{t,j-1})) = (RV^*{}_{t,j} - E\big(RV^*{}_{t,j}\big))^2. \tag{3.14}$$

There is no clear consensus on how to optimally evaluate intraday volatility. So we stick to the generally accepted and widely employed mean squared error that is similarly used in Engle and Sokalska (2012) and Behrendt and Schmidt (2018). Due to the enormous amount of data and following similar approaches in the literature, see e.g. Engle and Sokalska (2012), we restrict the window of our in-sample period to two months of data. The out-of-sample forecasts are computed based on a rolling window of two months from April to May 2017, that is the ARMA parameter are re-estimated for each single forecast based on the most recent data for two months.

Results appear to be similar to the daily forecast evaluation. There are stocks that can be forecasted more precisely with both cashtag or company name, but there are more stocks that do not show forecast improvements. This finding holds for Tweets with cashtags as well as for Tweets with company names. In fact, there is no clear advantage or disadvantage of either of the two specifications noticeable. Summary statistics presented in Table 3.8 show very similar patterns. Note that the change of the difference between the MSE of the respective Twitter model and the benchmark ARMA model in percent is displayed, such that numbers do not become too small. Once again, a negative sign indicates an improvement when incorporating Tweets. Regarding statistics from the full sample, we can see that the volatility of only approximately a quarter of the sample can be improved with the help of Twitter models. The median value of both specifications is positive while the mean is negative for cashtags models and positive for company names models, indicating that the benchmark model performs (slightly) better on average but that there are some potential outliers where cashtag models performance relatively well. The magnitude of improvement, however, is very small throughout the different specifications. On average, the MSE of cashtag models is 0.008% smaller than the MSE for the benchmark model and the MSE of company names models is 0.3% larger than the benchmark. In addition, they are not significantly different from zero for the vast majority of the sample, based on an (unreported) Clark-West test.

The sub-sample analysis reveals a very similar picture. The forecast performance of the majority of stocks cannot be improved with the help of Twitter models, although there are some stocks in the S&P100 where cashtag models are comparatively better performing. For SIFIs and stock indices, however, Twitter models on average over the whole sample, do not improve volatility forecasts.

Hence, we cannot conclude that Twitter information has any particular excess value over and above the standard volatilty models. One has to bear in mind that this result is also driven by the limited number of Tweets that are counted for most of the stocks. From Table 3.1 it can be seen that on average across the whole sample (based on the median), there is not even one Tweet per five minute interval. Of course, it could be intersting to further study why, how and when volatility forecasts of specific stocks can be improved. For the purpose of this exercise, however, we can conclude that intraday volatility models enhanced by Twitter data do not outperform standard models.

Table 3.8: Out-of-sample forecast evaluation ARMA-model

		Mean	Min.	1. Quan.	Median	3. Quan.	Max.
Overall							
	Improvement MSE Cashtag- BM	-0.0085	-5.8520	-0.0203	0.0078	0.0454	0.9442
	Improvement MSE Company name - BM	0.3062	-2.6713	-0.0201	0.0183	0.1232	16.9641
SP100							
	Improvement MSE Cashtag- BM	-0.0362	-5.8520	-0.0182	0.0007	0.0436	0.7951
	Improvement MSE Company name - BM	0.2133	-0.8376	-0.0254	0.0191	0.1706	4.5673
SIFI							
	Improvement MSE Cashtag- BM	0.0295	-0.4170	-0.0229	0.0073	0.0452	0.9442
	Improvement MSE Company name - BM	0.493248	-2.6714	-0.0066	0.0183	0.0968	16.9641
Index							
	Improvement MSE Cashtag- BM	0.19361	0.0295	0.0778	0.1259	0.2757	0.4253
	Improvement MSE Company name - BM	0.1067	-0.1109	-0.0599	-0.0090	0.2154	0.4398

This table shows the results of the out-of-sample intraday forecast analysis. The two models with Tweet component, cashtag and company name, are compared with the ARMA model. The evaluation is based on the mean squared error. Results are displayed in terms of the change of the respective criterion from the benchmark compared to the Twitter model for MSE in percent. Note, that positive number for these two criteria indicate superior performance of the benchmark model without Tweets, while negative values indicate better forecasts of Twitter models.

3.6 Conclusions

Data generated and disseminated online increasingly becomes valuable, also for research on financial markets. Being a popular web service for sharing condensed messages among a large user base, Twitter is suspected to contain valuable information for traders. This paper assesses the potential of activity on Twitter for improving short-run, i.e. from 5-minutes to one-day ahead forecasts of daily stock and index return volatilities compared to benchmark models. More than 150 stocks of systemically important banks, large international companies as well as several leading international stock indices are taken into account. The results are disillusioning. On average across the whole sample, there is hardly any significant forecast improvement when adding Twitter information. Only very rarely we can detect a significant improvement. Hence, based on our results we cannot recommend using Twitter information regarding volatility forecasts on a larger scale. However, it might be interesting to investigate why and how specific individual stock or index volatilities are indeed influenced by Twitter information. Twitter information is mostly noise with respect to financial markets. In some cases, however, volatility forecasts can be significantly improved with easy-to-obtain Twitter data and simple volatility models. More research is needed to identify patterns concerning which stocks are more impacted by Twitter activity and which ones less. In addition, filtering Twitter information could be useful to separate (financial market relevant-)signals from noise. But this comes at the expense of the general applicability of the identified approaches.

Furthermore, it could also be of interest to better understand the transmission channels from Twitter to investors' decisions. Given that more and more data will become available and will be employed due to the increasing digitalization of our societies, further research is required to elaborate further on the influence of, for instance, social media data on stock markets and stock market participants.

Contributions to the third essay "International cooperation on financial market regulation" by Michael Abendschein and Harry Gölz

The initial research design has been jointly developed by Michael Abendschein and Harry Gölz, including in particular the development of the model set-up. Both of them have also contributed to updates and revisions of the research design as well as drafting of the manuscript.

Harry Gölz has to a major extent carried out the mathematical analyses with a particular focus on propositions and proofs. He has also to a minor degree contributed to the interpretation of the achieved results within the framework of the research project.

Michael Abendschein has to a major extent interpreted the mathematical results with a view on the relation to financial regulation and cooperation. He has to a minor extent contributed to proofs and further mathematical analyses.

Michael Abendschein Harry Gölz

4 International cooperation on financial market regulation

4.1 Introduction

Modern financial markets are to an increasing extent characterized by cross-border linkages between banks from various countries. Growing capital flows or cross-border asset and debt positions as well as the ongoing emergence of multinational banking groups that operate across the globe have shed light on the necessity to coordinate regulatory oversight across different countries. As a consequence of the great financial crisis in 2007-09 policy makers and regulators have been forced to increase efforts to harmonize and coordinate common regulatory actions in order to strengthen the resilience of the global financial system. In practice, however, the number of bi- or multilateral supervisory agreements between national regulators is still limited, see Beck et al. (2019).

Our work is inspired by the need for a better understanding of regulatory cooperation. We characterize the formation process of self-enforcing supervisory cooperation and evaluate the feasibility and efficiency of cooperation. Our analysis provides novel insights into the difficulty of reaching a stable outcome by explicitly determining the number of countries being willing to cooperate. We show that in equilibrium a partial cooperative solution where only a fraction of countries jointly set regulatory standards and partly internalize the positive externalities from supervision can be stable.

We analyze endogenous coalition formation by applying a two-stage game from cartel theory developed by D'Aspremont et al. (1983). We prove for our baseline model that the number of countries forming a stable regulatory coalition is at most two. Furthermore, we highlight the crucial role additional club benefits might play and prove that larger coalitions are stable up to the grand coalition if they are taken into consideration; that is signatory countries of a coalition gain benefits over and above the joint welfare maximization. Such benefits can materialize in various forms, for instance through facilitated market access, more and faster information exchange and through reduced costs due to more adequate regulations adapted to the specific needs of the signatories. We contrast our findings choosing a different functional specification and find very similar results, highlighting the robustness of our results.

In line with the seminal paper from Dell'Ariccia and Marquez (2006) that studies cooperation and competition among regulators, we are also able to explain the economic rationale for regulatory cooperation given that individual welfare maximization on the country level leads to insufficiently low levels of contributions to global financial stability. In our model we can show

how the stability of the financial system, which is induced by regulations on the country level and operationalized in form of the so-called supervisory effort, can be increased when regulators coordinate their supervisory actions. Still, the difficulty of reaching stable coalitions is also highlighted and explained by free-riding incentives of individual countries.

In this work, we argue that efforts to coordinate joint regulatory policies resemble problems that have been put forward in the framework of public good analyses. The public good character of financial regulation has firstly been mentioned in White (1994) but has fallen into oblivion for a long time. More recently, Gaspar and Schinasi (2010), VanHoose (2016) or Agénor and Pereira da Silva (2018) have come up again with this idea, stressing the need to view and analyze supervisory cooperation by explicitly taking global externalities from national regulatory tools and efforts into account. Micro- and macroprudential tools such as minimum capital adequacy ratios or a deposit guarantee scheme imposed by a national regulator shall increase the stability of the respective domestic banking system through reduced incentives for risk-taking. Such efforts to sustain financial stability on the national level, however, translate into positive externalities within other countries that are interconnected via their respective financial system. Thus, there exists a social rationale to cooperate in safeguarding financial stability by national authorities which could lead to an increase of global financial stability, as brought forward for example by Dell'Ariccia and Marquez (2006) and Gaspar and Schinasi (2010).

On the contrary, however, incentives of national regulators to reduce risk-taking within their financial system is normally restricted by efforts to maintain or increase the competitiveness of their banking system, in particular when financial globalization is far reaching and domestic banks are deeply embedded in international competition. Higher risk leads to negative externalities in form of a destabilization of the international financial architecture, which may lead to banking failures or large economic downturns worldwide. Banks in countries with lax financial market regulation do not have to pay for taking higher risk positions and therefore obtain an implicit subsidy, as noted by White (1994). Hence, national regulators have to impose a certain level of supervisory effort that takes the trade-off between financial stability on the one hand and the banking system's performance on the other hand into account, see for example Dell'Ariccia and Marquez (2006).

In our view and in the work of Kaul et al. (1999) international financial stability is non-rivalrous and partly non-excludable and can be interpreted as an impure international public good. Thus,

if all countries that participate on global financial markets cooperate, social welfare in the involved countries will increase considerably. However, an increase of prudential supervision and regulation in a few countries leads to free-riding behavior of the remaining countries.

Early incentives to cooperate with respect to financial regulation were brought forward in the early 1970s when amounts of capital traded on euro markets increased drastically and the failure of the small German Herstatt Bank sent shock waves across the global financial markets, see Goodhart (2011). Consequently, several industrialized countries founded the Basel Committee on Banking Supervision (BCBS) which was meant to serve as a platform to discuss common regulatory incentives and reforms. Soon discussions started that aimed at creating a level playing field with respect to capital requirements induced by the Latin-American sovereign crises in 1982. Hence, the Basel Accord (Basel I) was a first important international financial regulation agreement (IFRA) with a view to reduce the implicit subsidy of domestic banking systems (Sinn 2003). The Basel Accord sets some voluntary standards that can only be understood as recommendations for all participating authorities. In recent work, Beck et al. (2019) collect data on supervisory cooperation worldwide and find more than 4000 agreements between individual countries. Their data reveals that the number of regulatory agreements is increasing but that only a small fraction of all potential agreements is realized.

Therefore, it is important to understand the drivers and incentives of countries to follow such voluntary regulations. Negotiations, for instance between American, European and Asian regulators, do regularly take place. They partly result in common binding standards while other issues remain disputed resulting in different regulatory approaches that could undermine the common goal of global financial stability. Beck et al. (2019) highlight the importance of costs and benefits from regulation. They also show that a large fraction of countries only cooperates with a small amount of other countries, hence the coalition size can often be assumed to be small.

We precisely model financial stability as an international impure public good within a multi-country model and assess the role cooperation might play to assure a stable global financial system. We endogenously determine the number of signatories by studying a game theoretical model of international cooperation that explicitly takes costs and benefits from regulation for each single country into account. A stylized fact of competition between national supervisors is incorporated in our model by assuming that costs from regulation on a national level increase

when other countries reduce their supervisory effort and vice versa, since a lower level of supervision abroad is directly interpreted as a competitive advantage compared to the more severely regulated national banking system. This setting allows us to assess how efficient the formation of an IFRA is in improving social welfare relative to the fully non-cooperative Nash equilibrium and the social optimal solution.

We show that partial cooperation leads to higher levels of supervisory effort on average and, hence, to a more stable financial architecture. Even though the signatory countries of the IFRA have to bear the additional regulatory efforts and costs, our results reveal that partial cooperation of two countries is stable for most of the feasible parameter space. Then, there is a distinct Pareto-improvement from partial cooperation relative to the non-cooperative Nash equilibrium. We furthermore confirm our results by studying the model under a different functional specification and find that coalitions of up to three countries might be feasible, irrespective of an additional assumption regarding parameter values. Our analysis also reveals that club benefits are of utmost importance when larger coalitions are desired. Whenever signatories of a regulatory coalition are exposed to additional benefits that go beyond the internalization of externalities, stable coalitions of any size can become feasible.

Two main implications can be drawn from these results. On the one hand, our model can confirm early empirical evidence that suggests that supervisory cooperation often materialize only within small groups of countries. This finding is particularly strong when gains from cooperation are small compared to involved costs (see Beck et al. (2019), and also for an analysis of the assumed underlying drivers). On the other hand, by confirming the necessity to cooperate with respect to regulatory actions, we can additionally show that the incorporation of club benefits is an important tool to overcome difficulties when aiming at the formation of larger stable supervisory coalitions.

The remainder of the paper is organized as follows: After the introduction of the research topic we discuss related literature in Chapter 4.2. In Chapter 4.3 we present the baseline model, explain the underlying assumptions and analyze the desirability of cooperation by comparing the non-cooperative Cournot-Nash equilibrium with the socially optimal solution. Subsequently, in Chapter 4.4 we examine the feasibility of IFRAs. First, we analyze the formation process of an IFRA with respect to an exogenously determined coalition size before we endogenously solve for the number of signatories. Then we add the notion of a club benefit and analyze its impact on coalition formation. As a robustness check, we solve the model with a different functional

specification in Chapter 4.5 before we discuss the implications of our results in Chapter 4.6. Concluding remarks will be presented in the final Chapter 4.7.

4.2 Related literature

The theoretical analysis of international financial stability plays an accretive important role in the economic literature. Allen and Gale (2004) describe the fundamental trade-off between financial stability on the one hand and the competitiveness of the banking industry on the other hand, similarly to Sinn (2003) who notices that national financial regulation creates a positive externality for foreign lenders, which leads to a low provision of regulation. Schoenmaker (2011) points to the same direction by interpreting financial stability, national financial policies and the global integration of financial markets as a trilemma whose aims cannot be jointly achieved.

While analyzing incentives to and outcomes of cooperation is relatively novel in the finance literature, it has a longstanding tradition in the area of public economics. VanHoose (2016) notes that similar problems of cooperation have already been extensively studied in the fields of fiscal policy, trade agreements or monetary policy and that potentially fruitful techniques and results developed so far are only slowly getting attention in the finance literature. Moreover, he indicates that the decision whether to cooperate in regulating financial markets appears to be very similar to research that has been conducted in the area of industrial organization for a long time. He argues that besides early work of White (1994), few efforts have been made to connect models and techniques from cartel theory to the obvious analogous of financial regulation cooperation.

This issue is most directly addressed in Gaspar and Schinasi (2010) who interpret financial regulation as a public good and suggest to rigorously apply models and techniques of the widely used toolbox of standard public economics. They apply the economic theory of alliances from Olson und Zeckhauser (1966) and the private provision of public goods model from Bergstrom et al. (1986) and compare the non-cooperative Nash equilibrium with the fully cooperative equilibrium. It turns out that the Nash equilibrium is characterized by an underprovision of financial stability since positive externalities of each individual regulator's effort to preserve financial stability is not accounted for in their utility functions. Additionally, they state that increasing cooperation is beneficial for all countries. They further introduce the notion of financial stability being an impure public good because benefits from regulation are assumed to

differ on a global and on a local level. They can show that regulatory cooperation could be established in such a framework and state that the intensity of cooperation depends on the share of public and country specific benefits. However, due to the considerable amount of restricting assumptions, they admit that the feasibility of such cooperation is unclear.

A more thorough analysis of the conflicting incentives a national regulatory authority faces is conducted in the seminal work of Dell'Ariccia and Marquez (2001, 2006).[20] Their model contains a representative bank per country, which is able to grant loans to domestic or international customers. Each bank is supervised by their domestic supervisor whose utility depends on the stability of the financial system on the one hand and the economic prosperity of its domestic bank on the other hand. Hardy and Nieto (2011) further develop the model from Dell'Ariccia and Marquez (2001) by explicitly incorporating a deposit guarantee scheme. Here, regulators can choose both a level of prudential supervision and the coverage of a deposit guarantee scheme. The latter decreases the costs that occur during a crisis while at the same time it increases the probability of a crisis by assuming incremental risk taking of banks. This feature is motivated by empirical work of Demirgüç-Kunt and Detragiache (2002) and Barth et al. (2006) who find excessive risk-taking by banks due to moral hazard when they can rely on large deposit guarantee schemes.

The set-up of Dell'Ariccia and Marquez (2006) and Hardy and Nieto (2011) allows to study the cooperation of national regulators. Dell'Ariccia and Marquez (2006) find that it is indeed hard to establish sustainable cooperation, in particular in a multi-country setting with heterogeneous countries. Even though the focus of their model is rather on banks' decision making and the financial cross border linkages and to a smaller extent on cooperation, they find that a supranational institution is more likely to be formed by homogeneous countries. In a two-country setting, their results show that capital requirements in a common regulatory regime must be higher than in each individual country in order to incentivize both regulators to join the cooperation. In this case, a two-country coalition is stable. However, in a multi-country setting, incentives to free-ride increase with the number of countries and they show that there are always countries that prefer to stay outside of an agreement. In the presence of a deposit guarantee scheme, the model of Hardy and Nieto (2011) shows in a two-country setting that cooperation

[20] An early, reduced version of their model is presented in Dell'Ariccia and Marquez (2001) which is used as a reference by some authors such as Hardy and Nieto (2011), while Dell'Ariccia and Marquez (2006) contains a more complete and more complex version of their model.

leads to higher levels of supervision and lower levels of deposit insurance compared to the non-cooperative case. They can show that a joint regulation with respect to prudential supervision and a national deposit guarantee scheme leads to a more efficient outcome compared to the fully non-cooperative setting but that additional cross-country cooperation with respect to deposit insurance could further enhance the efficiency of financial regulation. However, their work mostly remains silent with respect to the requirements or conditions necessary to establish and sustain such cooperation.

Eldridge et al. (2015) extend the model of Dell'Ariccia and Marquez (2006) by assuming two countries that only differ in their market size. They find that smaller countries have a greater incentive for lax supervision. Park and Kim (2018) also extend Dell'Ariccia and Marquez (2006) by formalizing the assumption that banks possess political influence. Within their setup, the likelihood of cooperation depends on the specific characteristic of regulatory effort being a strategic substitute or complement. Kara (2016) studies a two-country model with potentially asymmetric countries and finds that cooperation often increases aggregate welfare. The outcome of his model predicts on the one hand that cooperation among countries that are heterogeneous with respect to few characteristics is harder to establish than among homogenous countries. On the other hand, in case countries differ with respect to several characteristics, cooperation might be facilitated as well. There is little empirical work that covers aspects of regulatory cooperation with Beck et al. (2019) as a remarkable exception. They conduct a comprehensive study to evaluate the existence and stability of regulatory cooperation worldwide. They find that the intensity of cooperation is correlated with potential cooperation gains. These gains mainly stem from externalities which represent benefits and costs of cooperation.

Common to most of the research in this field is the finding that non-cooperation of national regulators leads to the provision of an inefficiently low level of regulation, such that the need for an international cooperation or the formation of a supranational regulatory institution could be supported. However, it remains unclear how to impose sustainable cooperation. Agénor and Pereira da Silva (2018) give an excellent overview on recent literature with respect to regulatory cooperation. They conclude that most papers suggest that cooperation increases social welfare, but that it is hard to quantify these gains correctly and to establish true cooperation in practice. More precisely, they criticize that most theoretical models only consider two countries while it might be harder to establish cooperation in the more realistic scenario with several countries.

While most of the research suggests that regulatory cooperation on the international level is

desirable with respect to global financial stability, some authors postulate a more differentiated view. Gehrig (2014) suspects a uniform global regulation to result in uniform bank business models that lead to increasing systemic risk. Based on the local information hypothesis which states that geographic location or distance is an important factor for banks' operations (see for instance Degryse and Ongena (2005) and Hauswald and Marquez (2006)), his model suggests that banks can choose whether to be regulated by a supranational or a local regulator. While global banks are better off by choosing the global regulator, local banks can profit from the superior information of their domestic regulator. While the problem could be solved by full information exchange between local and global regulators, Holthausen and Rønde (2005) show that such exchange is indeed hard to establish. More recently, Korinek (2017) presents a general equilibrium model that allows to determine if cooperation is beneficial. He emphasizes that cooperation is only advisable in case a non-cooperative equilibrium is Pareto-inferior to the cooperative equilibrium. Under a set of restricting assumptions, he argues that the non-cooperative Nash-equilibrium can be Pareto-efficient and potential spillovers turn out to be pecuniary externalities.

This work mostly relies on the work of Dell'Ariccia and Marquez (2006), Gaspar and Schinasi (2010) as well as Barrett (1994). We refrain from modelling bank's decision-making by studying the incentives of regulators to cooperate. We can show that the results of our model are in line with the predictions of Dell'Ariccia and Marquez (2006). However, we explicitly take the (impure) public good character of financial stability into account and rigorously apply techniques that have only been sketched by Gaspar and Schinasi (2010). This procedure allows us to model the decision of individual regulators to form or join a coalition endogenously. Our set up builds upon the canonical work with respect to international environmental agreements by Barrett (1994) and takes the impure public good modelling by Finus and Rübbelke (2013) as a starting point, before being adjusted in order to account for specific features of regulatory cooperation and competition.

This work contributes to the literature by studying the prerequisites of establishing and sustaining international regulatory cooperation and by explicitly examining the number of countries that are willing to form a cooperative coalition. We use straightforward modeling techniques that have proven to be useful in related fields but that have not been employed in this strand of the literature so far. Our work features some useful and realistic aspects that allow for a transparent interpretation of results against the background of empirical observations. Contrary to Hardy and Nieto (2011) we consider a multi-country setting in order to allow for a more realistic

interpretation of our results since regulatory cooperation is normally set up between several countries. By introducing countries' decision to sign an agreement as an additional stage of the game, our work also goes one step further than Dell'Ariccia and Marquez (2006). We explicitly model the formation of an agreement and are thus able to determine the number of countries within a cooperation endogenously. We can further show how the intensity of cross-border externalities influences the amount of countries being willing to cooperate. Our model is enhanced further with the novel introduction of additional club benefits of different magnitude. The effects of club benefits within the model seem to strengthen our view that their notion is necessary when thinking about regulatory cooperation. Consequently, our work allows to gain valuable insight on how to influence policies that are meant to increase global financial stability. This will also lead to a better understanding of empirical results that show that only smaller regulatory coalitions seem to be stable. Our work lies at the intersection of financial and public economics by applying tools from industrial organization and environmental economics to a current topic with respect to financial regulation.

4.3 The model of financial stability as an impure public good

4.3.1 Assumptions

We investigate the incentives of countries to regulate their domestic banks by analyzing countries' benefits and costs of financial regulation. We consider a world of $N \geq 2$ ex-ante symmetric countries, indexed $i = 1, \dots N$. All countries are interconnected on global financial markets. By assumption all countries have an incentive to sustain financial stability and thus supply supervisory effort. Like Hardy and Nieto (2011), this is our choice variable in the model. Alternatively, supervisory effort could be interpreted more directly as for example the capital adequacy ratio that is set by the regulator, see e.g. Dell'Ariccia and Marquez (2006). However, Agénor and Pereira da Silva (2018) note that it is difficult to ascribe the consequences of supervision to a distinct supervisory tool since there are always various tools simultaneously in use, be it micro-, macroprudential, or even monetary policy instruments. Hence, we stick to the rather general notion of supervisory effort. [21]

We assume that financial market supervision constitutes an impure public good, as outlined in

[21] See also Acharya (2003) and Buck and Schliephake (2013) for an analysis regarding the interaction of different regulatory tools within the framework of a theoretical model.

the following. Contrary to Dell'Ariccia and Marquez (2006) and in line with Dell'Ariccia and Marquez (2001) we take a reduced form approach and we abstract from the concrete loan and monitoring decision of banks. Instead of a rather general functional form as in Dell'Ariccia and Marquez (2001), we use a parametric model in order to be able to derive more specific results.

Each country i obtains a benefit $B_i(\bar{q})$ from the supervision of all international financially connected countries, where $\bar{q} = \frac{1}{N}\sum_i^N q_i$ denotes average supervisory effort. We assume a quadratic public benefit function which is given by:

$$B_i(\bar{q}) = h\bar{q} - \frac{1}{2}\bar{q}^2, \tag{4.1}$$

where h is a positive parameter. Thus, each country receives an equal share of worldwide supervisory effort which can be thought of e.g., as the implementation of capital requirements that are supposed to increase the stability of a single financial institution or, more generally, as the degree of prudence in the exercise of supervisory power. The reasoning is straightforward because a positive relation between supervisory effort and global stability is assumed. The more effort is undertaken in global terms, the more unlikely it becomes to experience financial crises. This translates into shrinking expected negative implications for each individual country. A more stable financial system overall decreases the possibility for each bank and, hence, each country to be confronted with sudden bank failures that are caused by financial distress somewhere else.

Moreover, countries earn an extra exclusive benefit $D_i(q_i)$ from their individual supervisory effort which is given by:

$$D_i(q_i) = hq_i - \frac{1}{2}{q_i}^2. \tag{4.2}$$

It is suggestive that regulation on a national level increases in particular the stability of the respective country over and above the contribution of the overall system's stability for two main reasons. First, and following the literature (see e.g. Dell'Ariccia and Marquez (2006)), we assume home country regulation, which implies that a national regulatory authority oversees its own domestic banks only.[22] One could argue that this assumption is not fully met with respect to well-known regulatory cooperation such as the Euro zone. However, we aim to set up a

[22] See Park and Kim (2018) who assume host country regulation as an exception.

general model without referring to a specific regulatory set-up. Implications for our results in an alternative host country regulatory environment are discussed in Chapter 6. Given that domestic banks can be expected to have a larger market share in their domestic market, it can be assumed that the domestic financial system is particularly stable. This view is supported by the argument put forward in Gosh et al. (2017) who find that sound supervision on a national level reduces the risk of a country's financial system to be hit by shocks that stem from foreign financial shocks. Yet financial stability is often merely understood in global terms, there is still a reason to assume additional regional benefits.

As our model incorporates many features of the seminal model of regulatory cooperation presented in Dell'Ariccia and Marquez (2006), we will illustrate how our set-up builds upon their framework in the following. In our model, the regulator's welfare is on the one hand affected by the stability of the financial system. It is directly influenced by a supervisor's own strength of supervisory effort as well as the cumulative effort of the other supervisors. In the framework of their partial equilibrium model, Dell'Ariccia and Marquez (2006) similarly assume that financial stability constitutes one of the two drivers of a regulator's utility. Financial stability materializes by the monitoring effort that banks exert on behalf of their loan portfolio, which is directly influenced by capital standards set by their respective supervisor. More precisely, the utility of individual regulators depends on the stability of the financial system directly formalized by the (scaled) regulatory effort and on the other hand on the profitability of the banking system expressed by its cumulative profits achieved from lending activities. In our model, financial stability arises from supervisory efforts on the national level which enter our benefit function in equations (4.1) and (4.2) as well as on the efforts of other regulators additionally captured by equation (4.1). This results in a concave benefit function, similar to Dell'Ariccia and Marquez (2001).

The benefits stemming from bank profitability in Dell'Ariccia and Marquez (2006) are interpreted as costs of supervision in our model, because supervisory efforts directly lead to shrinking bank profits. Hence, instead of taking bank profits directly into account, we interpret the shrinking profitability as costs directly linked to supervisory effort that enter the payoff function of our regulator with a negative sign.

The downside of supervising domestic banks $C_i(q_i, \bar{q})$ depends, in our model, on two elements.

$$C_i(q_i, \bar{q}) = \frac{1}{2} c q_i^2 + \theta(q_i - \bar{q}), \tag{4.3}$$

where $c > 0$ and $\theta > 0$. The first part covers the direct costs of individual supervisory efforts that are modelled in a straightforward manner by a quadratic function. It captures the decline of banks' profitability due to costly regulatory constraints on the domestic level and thus only depends on individual effort. Here, our model mirrors the effects of the partial equilibrium model in Dell'Ariccia and Marquez (2006) where regulators set a specific capital adequacy ratio. The introduction or increase of a capital ratio decreases the amount of loans given by each bank and eventually decreases their profits.[23]

The second expression reflects the change in competitiveness of domestic banks. We assume that bank loans are imperfect substitutes for the consumers. Once again, an increase in the degree of supervisory effort on a domestic level will reduce the competitiveness of domestic banks when facing competition with less restricted foreign banks. A higher (lower) than average supervisory effort implies a loss (gain) in competitiveness of domestic banks. Here, we account for an effect similarly observable in Dell'Ariccia and Marquez (2006). In their model, the profitability of the domestic banking system does not only depend on the capital ratio set by their own domestic regulator, but also on the stringency of foreign regulators and, hence, on domestic bank's competitiveness with respect to foreign competitors. (Recall that by assumption, the capital requirements set by regulators can only be imposed upon their own domestic banks). Banks face a loan demand curve where consumers can choose between loans from domestic and foreign banks. The higher the capital ratio imposed upon domestic banks, the larger the share of loans originated by foreign banks on the domestic loan market. In our model, the second part of the cost function captures this competitiveness effect by assuming an additional benefit in case domestic regulation is weaker than the average abroad (reflecting the increasing profitability of the domestic banking sector abroad) and vice versa by assuming additional costs in case domestic supervisory effort is larger than in foreign markets.

To sum up, the assumed net benefit function of country i is given by:

$$\pi_i = \alpha\left(h\bar{q} - \frac{1}{2}\bar{q}^2\right) + (1-\alpha)\left(hq_i - \frac{1}{2}{q_i}^2\right) - \frac{1}{2}c{q_i}^2 - \theta(q_i - \bar{q}), \tag{4.4}$$

[23] See De Jonghe et al. (2020) and Uluc and Wieladek (2017) for a recent discussion on the relationship between supervisory effort and bank lending.

where $\alpha \in (0,1)$ is the public good factor of financial stability. If $\alpha = 1$, financial stability would be a pure public good and if $\alpha = 0$, it would be a private good. The functional specification we use here has also similarly been used by Barrett (1994) and Finus and Rübbelke (2013) in the literature on environmental treaty formation. The net benefit functions of all countries are assumed to be common knowledge and supervisory effort levels are freely observable by each country.

Finally, note that α and θ can be interpreted as proxies for the strength of financial cross-border activities and likewise as proxies for financial interconnectedness. The larger α, the stronger a country depends on the efforts of foreign regulators to sustain financial stability because the own country depends to a great extent on global financial stability. This is often the case when countries are strongly active and interconnected on global financial markets. θ eventually expresses the degree of competition a country's banking system is exposed to regarding foreign banks. The more open financial markets are, the more (foreign) banks operate in a country and the stronger the degree of competitiveness shall be assumed. Thus, the stronger cross-border activities in the financial sector are assumed, the larger will be the effect of different regulatory regimes at home and abroad which translates into larger values for θ.

4.3.2 Cournot-Nash equilibrium

In the non-cooperative Cournot-Nash game all countries choose supervisory effort simultaneously taking the effort levels of all others as given. In this case every single country simply maximizes its individual net benefit of supervisory effort and does not internalize the positive externalities. The maximization problem of each country i is given by:

$$\max_{q_i} \pi_i = \alpha\left(h\bar{q} - \frac{1}{2}\bar{q}^2\right) + (1-\alpha)\left(hq_i - \frac{1}{2}{q_i}^2\right) - \frac{1}{2}c{q_i}^2 - \theta(q_i - \bar{q}) \qquad (4.5)$$

The corresponding first order condition is given by:

$$\alpha\left(\frac{h}{N} - \frac{\bar{q}}{N}\right) + (1-\alpha)(h - q_i) - cq_i - \theta\left(1 - \frac{1}{N}\right) = 0. \qquad (4.6)$$

We solve equation (4.6) by using $\sum_i^N q_i = q_i + Q_{-i}$, where $Q_{-i} = \sum_{j \neq i} q_j$ in order to compute the best response function of country i, which is given by:

$$q_i(Q_{-i}) = \frac{N(h(N+\alpha-N\alpha)+\theta-N\theta)-\alpha Q_{-i}}{N^2(1+c-\alpha)+\alpha}. \tag{4.7}$$

$\partial q_i/\partial Q_{-i} < 0$, which implies that the optimal effort choice of each country negatively depends on the aggregate effort level of all other countries. Concretely, this implies that a higher regulatory effort of one country is crowded out by the free-riding behavior of the others. Thus, effort levels are strategic substitutes. To compute the Cournot-Nash equilibrium we use the symmetry of the countries and solve equation (4.6) by inserting $\sum_i^N q_i = Nq_i$. Individual and average effort quantities in equilibrium are given by:

$$q_u = \bar{q}_u = \frac{h(N+\alpha-N\alpha)+\theta-N\theta}{N(1+c-\alpha)+\alpha}, \tag{4.8}$$

where index i is replaced by u indicating the non-cooperative Cournot-Nash solution. The partial derivatives are $\partial q_u/\partial h > 0$ and $\partial q_u/\partial c,\ \partial q_u/\partial \alpha, \partial q_u/\partial \theta, \partial q_u/\partial N < 0$. It is important to note that regulatory incentives decrease with the public good factor of financial stability α and the competitiveness factor θ. Furthermore, the last partial derivative indicates the familiar finding that free-riding incentives increase with the amount of countries interconnected.

We summarize our main findings in Proposition 1 where the proof is given in the Technical Appendix C.

Proposition 1: For the Cournot-Nash equilibrium of the model specification given by equation (4.4), the following implications hold:

1.1 *The more important the public good character of financial stability α, the less willing is a country to contribute to regulation ($\partial q_u/\partial \alpha$).*

1.2 *The more important the competitiveness θ, the less willing is a country to contribute to regulation ($\partial q_u/\partial \theta$).*

1.3 *Free-riding incentives increase with the number of countries N interconnected on global financial markets ($\partial q_u/\partial N$).*

4.3.3 Social optimum

In order to evaluate the Cournot-Nash solution, we compare it to a socially desirable solution where all countries cooperate and hence all externalities are internalized. In the full cooperative

outcome countries maximize aggregate net benefits of supervisory effort and solve the following maximization problem:

$$\max_{q_i} \pi_i = \alpha\left(h\frac{Nq_i}{N} - \frac{1}{2}\left(\frac{Nq_i}{N}\right)^2\right) + (1-\alpha)\left(hq_i - \frac{1}{2}{q_i}^2\right) - \frac{1}{2}c{q_i}^2 - \theta\left(q_i - \frac{Nq_i}{N}\right). \tag{4.9}$$

The maximization problem in equation (4.9) is equivalent to the maximization problem where $\max_{q_i} \sum_{i=1}^{N} \pi_i$. Due to symmetric concave net benefits only a symmetric equilibrium is optimal. Here, the maximization is conducted jointly across all countries and it is clear ex-ante that the average regulatory effort is identical for each country. The corresponding first order condition is given by:

$$\alpha(h - q_i) + (1-\alpha)(h - q_i) - cq_i = 0 \tag{4.10}$$

and individual as well as average Pareto-efficient effort levels are given by:

$$q_o = \bar{q}_o = \frac{h}{1+c}, \tag{4.11}$$

where index i is replaced by o in order to indicate the socially optimal solution.

Using the results above, we prove in the Technical Appendix C that $q_o > q_u$. Thus, countries are better off if they fully cooperate instead of choosing effort levels non-cooperatively. Furthermore, the degree of inefficiency determines the desirability of cooperation. We measure this with the difference $q_o - q_u$. Hence, cooperation is more beneficial the larger the effort gap between the Nash equilibrium and the social optimum. We show in the Technical Appendix C that this gap increases with the intensity of both externalities. They eventually create the rationale for cooperation in this model.

The economic intuition for this finding is clear. The parameters α and θ represent cross-border dependence as they measure the influence other countries exert on a single country. The larger α, the more a single country depends on supervisory efforts on a global level. In such circumstances, financial stability is nearly a pure public good and hence the necessity to cooperate is evident. The case for θ is closely related. The larger it is, the stronger the competitiveness of domestic banks compared to foreign banks influences the utility function of a single country. In order to circumvent costly competition, it becomes more necessary to coordinate common regulatory actions that influence each country's bank profitability.

We summarize our main findings in Proposition 2.

Proposition 2: Comparing the Cournot-Nash equilibrium with the social optimum, the following implications hold:

2.1 *Supervisory effort levels in the Cournot-Nash equilibrium are inefficiently low* $(q_u < q_o)$.

2.2 *The desirability of cooperation* $(q_o - q_u)$ *increases with the intensity of cross-border externalities* α *and* θ.

Proposition 2.1 confirms Proposition 1 of Dell'Ariccia and Marquez (2006). Proposition 2.2 is also in line with Dell'Ariccia and Marquez (2006) and additionally with the empirical findings of Beck et al. (2019). In what follows, we analyze a model version where partial cooperation is a possible outcome, which is in line with real world observations where cooperation is normally realized within a sub-group of countries.

4.4 The model of an international financial regulatory agreement

The model variant in Chapter 4.3 assumes that countries simply choose their effort levels and behave either non-cooperatively or fully cooperatively. In this chapter, we extend the basic model in three dimensions. In Chapter 4.4.1, we first modify the model by assuming that a cooperative coalition of given size plays Cournot-Nash against the free riders. In Chapter 4.4.2, we additionally endogenize the coalition size by assuming that countries are able to decide whether they want to become a signatory of a coalition or not. We further assume in Chapter 4.4.3 that signatories additionally gain exclusive club benefits.

4.4.1 Partial cooperation of given size

In this chapter we still assume a simultaneous game, but now an exogenously given coalition acts cooperatively by fully internalizing the external effects within the group. Hence, signatories maximize their joint net benefits whereas free riders maximize their own net benefits as before. We assume that there exist an exogenous number of $n \in \{1, \dots N\}$ signatories and $N - n$ free riders. If $n = 1$ the coalitional Nash equilibrium coincides with the ordinary Nash equilibrium in Chapter 4.3.2. If $n \geq 2$ a cooperating coalition is formed and the equilibria differ. A signatory country is denoted by subscript *s* and a free rider country is denoted by subscript *f*. The free riders take the effort levels of all other countries as given and maximize their own net

benefit by choosing the effort level unilaterally. The maximization problem of free riders is given by:

$$\max_{q_f} \pi_f = \alpha\left(h\bar{q} - \tfrac{1}{2}\bar{q}^2\right) + (1-\alpha)\left(hq_f - \tfrac{1}{2}{q_f}^2\right) - \tfrac{1}{2}c{q_f}^2 - \theta\left(q_f - \bar{q}\right). \qquad (4.12)$$

The corresponding first order condition is given by:

$$\alpha\left(\frac{h}{N} - \frac{\bar{q}}{N}\right) + (1-\alpha)\left(h - q_f\right) - cq_f - \theta\left(1 - \frac{1}{N}\right) = 0. \qquad (4.13)$$

We solve equation (4.13) by using $\sum_i^N q_i = Q_f + Q_s$, where $Q_f = (N-n)q_f$ and $Q_s = nq_s$, in order to compute the aggregate best response function of free riders, which is given by:

$$Q_f(Q_s) = \frac{(N-n)(hN(N-(N-1)\alpha) - (N-1)N\theta - \alpha Q_s)}{N(N(1+c-\alpha)+\alpha) - n\alpha}. \qquad (4.14)$$

$\partial Q_f / \partial Q_s < 0$ implies that the optimal effort choice of free riders negatively depends on the aggregate effort level of the coalition.

The signatories of the coalition take the effort levels of all free riders as given and maximize their joint net benefits such that:

$$\max_{q_s} \pi_s = \alpha\left(h\left(\frac{nq_s + (N-n)q_f}{N}\right) - \frac{1}{2}\left(\frac{nq_s + (N-n)q_f}{N}\right)^2\right) \qquad (4.15)$$
$$+ (1-\alpha)\left(hq_s - \frac{1}{2}{q_s}^2\right) - \frac{1}{2}c{q_s}^2 - \theta\left(q_s - \frac{nq_s + (N-n)q_f}{N}\right)$$
$$.$$

The first order condition is given by:

$$\theta(\tfrac{n}{N} - 1) + (1-\alpha)(h - q_s) - cq_s + \alpha(\tfrac{hn}{N} - \tfrac{n(nq_s + (N-n)q_f)}{N^2}) = 0. \qquad (4.16)$$

We solve equation (4.16) by using $q_s = \frac{Q_s}{n}$, $q_f = \frac{Q_f}{N-n}$ and solve for Q_s in order to get the aggregate best response of the signatories which is given by:

$$Q_s(Q_f) = \frac{n(hN(N + n\alpha - N\alpha) + (n-N)N\theta - n\alpha Q_f)}{N^2(1+c-\alpha) + n^2\alpha}. \qquad (4.17)$$

Inserting (4.14) into (4.17) we get the total optimal effort level of the signatories Q_s^* given by:

$$Q_s^* = \frac{hn\Big((n-1)n\alpha - (N(\alpha-1) - n\alpha)\big((1+c)N - (n+N-1)\alpha\big)\Big) - n(n-N)(-(1+c)N + (n+N-1)\alpha)\theta}{(1+c-\alpha)((1+c)N^2 + (n-N)(n+N-1)\alpha)}, \tag{4.18}$$

where individual effort levels are $q_s^* = \frac{Q_s^*}{n}$. Inserting (4.18) into (4.14) we get the optimal aggregate effort level of the free riders given by:

$$Q_f^* = \frac{(n-N)(hN^2(1+c-\alpha)(\alpha-1) + h(n-1)n(\alpha-1)\alpha + hN\alpha(\alpha-c-1) + (N-1)N(1+c-\alpha)\theta + (n-1)n\alpha\theta)}{(1+c-\alpha)((1+c)N^2 + (n-N)(n+N-1)\alpha)}. \tag{4.19}$$

Individual efforts of free riders are $q_f^* = \frac{Q_f^*}{N-n}$ and average effort levels are $\bar{q} = \frac{Q_s^* + Q_f^*}{N}$. By inserting all equilibrium effort levels into the payoff functions we receive the equilibrium payoffs of the signatories π_s and π_f.

We summarize our findings in Proposition 3 and the proof is given in the Technical Appendix C.

Proposition 3: For the coalition-fringe equilibria $(n \geq 2)$ *of the model specification given by equation (4.4), the following implications hold:*

3.1 *Supervisory effort levels of signatories* q_s^* *are higher and the effort levels of free riders* q_f^* *are lower than in the Cournot-Nash equilibrium. Average effort levels are higher than in the Cournot-Nash equilibrium.*

3.2 *The effort levels of free riders* q_f^* *are strictly decreasing in n and the effort levels of signatories* q_s^* *are increasing in n if* $\alpha \leq \frac{1}{2}$ or $c \geq 1$.

3.3 *The net benefits of free riders* π_f *are strictly increasing in n and the net benefits of signatories* π_s *are increasing in n for* $n > \sqrt{N}$.

3.4 *The net benefits of free riders* π_f *are strictly higher than the net benefits of signatories* π_s.

From Proposition 3.1 we note that the higher efforts of the cooperating coalition are just partly crowded out by the lower efforts of the free-riding countries. Consequently, cooperative solutions increase aggregate financial stability even though free riders reduce their own efforts given a coalition of signatories.

Proposition 3.2 shows that free-riding incentives of countries remaining outside the coalition increase when the intensity of cooperation increases. The more efforts are made by cooperating countries, the more financial stability is ensured, so that free rider countries increasingly have to implement less costly regulation of their own. On the contrary, if countries decide to cooperate they are in most cases rewarded by higher efforts of the other signatories. More precisely, that happens when c is at least one or α is at most 0.5. Whenever costs of regulation are high, the effort levels of signatories depend more on the number of signatories as countries are only willing to invest in costly regulation if they know that others do so, too. If regulation is cheap, incentives to conduct regulations are to a smaller extent influenced by other signatories. Similarly, whenever α is small and financial stability is rather a private than a public good, incentives to cooperate are weaker such that more regulatory efforts are only made if the other countries mirror this behavior as a country's intrinsic motivation to cooperate is limited.

Proposition 3.3 illustrates that signatories are in general better off when more countries cooperate as this implies a larger coordinated action to provide welfare increasing regulation. However, solutions exist for small n when joining an agreement does not pay off. Simultaneously, free riders are always better off with an increasing number of signatories as this enlarges potential benefits from setting their own lower standards. Based on the given restricting assumptions, the net benefits of both types are increasing with the intensity of cooperation, which generally induces an incentive to cooperate. However, Proposition 3.4 reveals that there exists an incentive to stay outside the coalition because net benefits are strictly higher by staying outside.

These results illustrate countries' incentives for cooperation because with an exogenously given number of signatory countries all countries are better off with cooperation. However, they also reveal a potential difficulty to reach intensive cooperation, since incentives to free ride on the coalition's higher efforts exist. In the following chapter, we jointly analyze incentives to sign and to defect from a coalition by analyzing the stability of an agreement.

4.4.2 Endogenous coalition formation

We extend the model of Chapter 4.4.1 by assuming that countries have two decisions. First, they can choose to join a coalition, second, they choose their respective effort level. The additional choice to sign a cooperative agreement allows us to endogenize n. We concretely analyze the IFRA formation using a two-stage non-cooperative game theoretical model which originates

from the theory of industrial organization and is regularly applied in the literature on international environmental agreements. This model is well suited to analyze the incentives for cooperation in the absence of a supranational institution that is able to enforce binding regulatory measures. Hence, sovereign countries voluntarily sign a treaty to safeguard financial stability. Concretely, the IFRA formation consists of the following two stages:

Stage 1: All countries simultaneously decide whether they want to become a signatory of the IFRA or not.

Stage 2: Cooperative signatories and non-cooperative free riders choose their effort levels simultaneously.

Stage 2 can be regarded as the effort choice subgame, which assumes a Cournot-Nash structure as in Chapter 4.4.1.[24] The entire game is solved via backward induction and the solution denotes a subgame perfect equilibrium.

We solve the game starting with the second stage, where free riders maximize their own net benefits and signatories maximize their joint net benefits taking the decisions with respect to n as given. The solution of the subgame in stage 2 is equal to the solution of the model in Chapter 4.4.1. The difference, however, is that the outcomes of the effort subgame are the equilibrium payoffs of the signatories $\pi_s(n)$ and of the free riders $\pi_f(n)$ which depend on the endogenous variable n. The payoffs are needed in order to solve the first stage of the model. The solution concept, which is used to solve this stage, is based on D'Aspremont et al. (1983) and originates from the literature on cartel stability. According to this concept, an IFRA consisting of n signatories is stable if:

$$\pi_s(n) \geq \pi_f(n-1) \quad \text{and} \quad \pi_f(n) \geq \pi_s(n+1), \tag{4.20}$$

where the first expression denotes the internal stability and the second expression the external stability condition. Internal stability means that no signatory country has incentives to withdraw from the agreement because the benefits within a cooperating coalition of given size n, $\pi_s(n)$ are larger or equal to the benefits when leaving the coalition and facing benefits $\pi_f(n-1)$ according to the diminished number of countries within the coalition, $n-1$. Similarly, external stability requires that free riders have no incentive to join the IFRA by assuming that all others

[24] In the literature on environmental agreements the Cournot-Nash variant is used by Finus and Rübbelke (2013), whereas Barrett (1994) assumes Stackelberg-leadership of the cooperating coalition.

stick to their decisions given that their benefits were reduced when entering the coalition. The equilibrium is achieved because the participation of one additional country would increase the total number of signatories and also signatories' level of effort. On the contrary, if a signatory withdraws from the IFRA, the remaining ones decrease their effort levels. This mechanism can be interpreted as a system of credible penalties and rewards. By applying the stability conditions, we are able to determine the absolute degree of cooperation. However, a small coalition can potentially be quite efficient when signatories increase their effort levels substantially. Therefore, we need a measure that is able to value the intensity of the welfare improvement of the partial agreement.

In order to assess the efficiency of the agreement further, we need to compare its performance with the social optimal and the non-cooperative outcome. Consequently, we use as a relative welfare measure, the "closing the gap index" (CGI), which was developed by Eyckmans and Finus (2006). The CGI $\in [0,1]$ and measures the welfare enhancement due to the IFRA formation relative to the gap between the socially desirable situation and the situation where none of the countries cooperates. The total net benefit in case of treaty formation is given by $\Pi_a = n\pi_s(n) + (N-n)\pi_f(n)$. We get total payoffs in the social optimum Π_o by inserting $n = N$ in Π_a and we get total non-cooperative payoffs Π_u by inserting $n = 0$. The CGI is thus given by:

$$CGI = \frac{\Pi_a - \Pi_u}{\Pi_o - \Pi_u}. \tag{4.21}$$

When $CGI = 1$, the IFRA is most efficient and is able to close the gap entirely. On the contrary, when $CGI = 0$, the IFRA is fully inefficient and does not improve the situation at all.

We summarize our findings for this model variant in Proposition 4, with the proof given in the Technical Appendix C.

Proposition 4: For the self-enforcing agreement of the model specification given by equation (4.4), the following implications hold:

4.1 *Two countries always cooperate in equilibrium when $N = 2$. Two countries cooperate if $c \geq -1 + \frac{(N^2+N-4)\alpha}{N^2} + \frac{2\alpha\sqrt{3+(N-3)N}}{N^2}$ and otherwise the non-cooperative case is stable when $N \geq 3$.*

4.2 *Incentives to sign an agreement do not increase with the intensity of cross-border externalities α and θ.*

4.3 The CGI is invariant in the benefit parameter h and in the competitiveness factor θ.

4.4 The CGI converges to zero, if the number of countries N becomes large.

Proposition 4.1 confirms the finding of Dell'Ariccia and Marquez (2006) indicating that the grand coalition of two symmetric countries is always stable if $N = 2$. Moreover, we show that in the case of $N \geq 3$ a coalition larger than two can never be stable while a two-country coalition is stable for the parameter space specified above. A coalition of two countries is unstable only in the case of extreme parameter constellations with α being very large and N and c being very small. In particular in a situation with very low costs of supervision (c very low) and a high degree of free-riding (α very large), entering a two-country coalition decreases welfare. As we have argued above, financial stability is clearly an impure public good, so α is likely to be not too close to one. In this case, each country supplies a considerable amount of supervisory efforts because it gains from financial stability without having to pay too much. Hence, coordinated actions to increase supervisory efforts are neither necessary nor worthwhile. This scenario, however, does not appear to be very plausible as costs to regulation can be assumed to be considerably different from zero. Proposition 4.1 is also related to Propositions 5 and 6 of Dell'Ariccia and Marquez (2006), but our results are more specific. We find that there always exists an incentive not to sign the grand coalition if $N > 2$. We additionally note that a stable two-country coalition can break down when there are financial links with further countries, which complements Proposition 6 in their work.

From Proposition 4.2 we note that incentives to sign an agreement do not depend on the degree of excludability of financial stability. This result is in line with Finus and Rübbelke (2013), who find that additional private benefits of climate protection do not increase the incentive to sign a climate treaty. For our model, this implies that the private benefit on the national level stemming from financial regulation does not create further incentives to cooperate. Even though cooperation becomes more beneficial with higher α, incentives to free-ride also increase. Incentives for cooperation do not depend on θ either. On the one hand, θ could have been expected to be a driver of cooperation as each country can be assured not to provide more regulatory effort than the other countries when all countries set a common level. On the other hand, θ reinforces incentives to deviate as it pays-off to undercut a common effort level. In our model, these two effects cancel each other out. Proposition 4.3 shows that the relative efficiency of an

agreement of two countries does not depend on the competitiveness externality nor on the specific parametrization of the benefit function. Accordingly, the intensity of public benefits α, the direct costs c and the group size N are the drivers of relative efficiency. Proposition 4.4 illustrates that the positive effect of a two-country agreement vanishes if N is very large. This is a natural result as the effect of efforts to maintain global financial stability are marginal if only two countries jointly set common effort levels while simultaneously a very large number of countries free-rides. Still, this result highlights the suitability of our model in explaining characteristics of regulatory agreements.

In order to illustrate the solution procedure and the results of the membership stage we use a numerical example where the parameters are $c = 2$, $\alpha = 0.7$, $h = 50$, $\theta = 1$ and $N = 10$. Figure 4.1 depicts the stability functions for this numerical example.

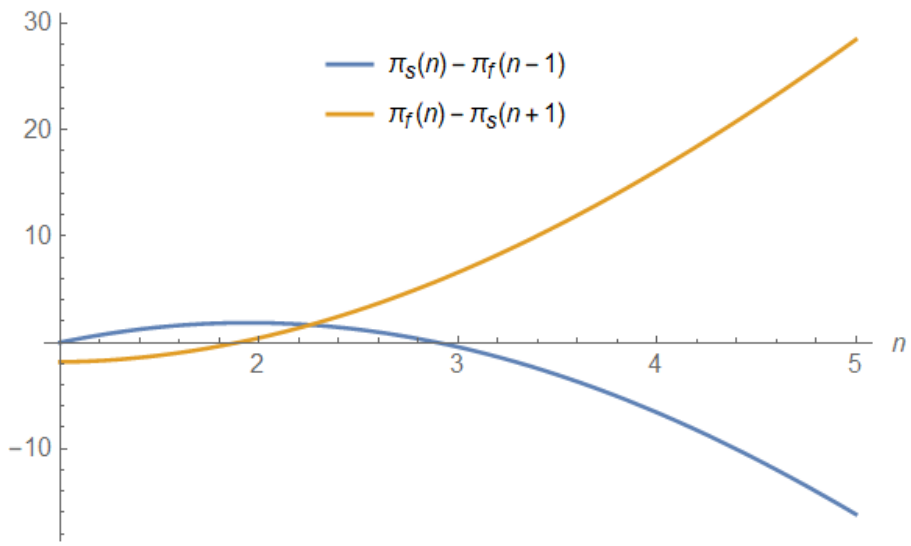


Figure 4.1:Stability functions: numerical example.

The stability functions for internal and external stability are displayed for the given numerical example. The benefit of signatories and free riders depends on the number of signatories n. There is only one interval where both functions take on positive values and, hence, n=2 is the stable coalition size of this simulation.

We see that there exists only one interval where both stability functions take on positive values and this interval only contains one integer which is $n = 2$. Moreover, we further illustrate the numerical example in Table 4.1. The decision to join the coalition in this model is a one-shot decision, but it seems to be useful to think of it as a sequential decision. We first analyze the external stability condition beginning in the case where $n = 1$ and none of the countries cooperates. A country decides to join the coalition if the following payoff is higher than staying

outside of the agreement. From Table 4.1 we see that $\pi_f(1) = 288.6 < 290.4 = \pi_s(2)$. Next, we assume that there are already two countries which have signed the IFRA. In this case $\pi_f(2) = 296.5 > 296.1 = \pi_s(3)$ and, hence, there does not exist an incentive to enter a coalition of size three. Thus, the external stability is fulfilled for $n \geq 2$. Subsequently, we check when the internal stability condition holds and conduct the analysis in a situation where all countries fully cooperate. If the first country defects, it will get $\pi_f(9) = 506.1 > 416.7 = \pi_s(10)$ and the country decides to withdraw. It pays off for countries to withdraw whenever $n > 2$. Thus, the only stable situation where both stability conditions are fulfilled is the case where two countries sign the IFRA.

Table 4.1: Internal and external stability: numerical example

n	$\mathbf{q_s}$	$\mathbf{q_f}$	$\boldsymbol{\pi_s}$	$\boldsymbol{\pi_f}$
1	7.43	7.43	288.6	288.6
2	8.75	7.42	290.4	296.5
3	10.03	7.4	296.1	312.1
4	11.26	7.38	305.5	334.3
5	12.41	7.35	318.2	362.3
6	13.47	7.32	333.8	394.7
7	14.43	7.28	352	430.3
8	15.28	7.24	372.2	467.8
9	16.03	7.19	393.9	506.1
10	16.67	-	416.7	-

The effort levels q and resulting benefits π from signatories and free riders conditional on the number of signatories n are displayed. The only situation where the benefit of a free rider π_f does not increase when it joins the coalition and simultaneously the benefit of signatories π_s does not increase when it leaves the coalition is when $n = 2$.

The CGI for the numerical example is 0.053. This implies that the agreement is able to close 5.3% of the gap between the fully non-cooperative and the social optimal case.

Furthermore, Figure 4.2 shows the CGI depending on the cost parameter and the number of countries when $\alpha = 0.7$. The CGI obviously decreases in c and N. If N is very small the agreement consisting of two countries is relatively efficient. In case c is close to zero, supervisory effort is rather cheap and the gap between the non-cooperative and the full cooperative case is very small. Figure 4.3 illustrates the CGI depending on the public good parameter and the number of countries when $c = 2$. The CGI is slightly increasing in α even though the absolute

degree of cooperation is constantly $n = 2$. Hence, a cooperative agreement is most efficient when it is needed the most.

So far, our impure-public-good-approach is not fully able to explain all existing cooperative solutions regarding financial regulation, as it cannot give a rationale for multilateral cooperative arrangements such as the Basel Accords where many countries agree upon common standards. Henceforth, we extend our stylized model in the following section by taking into account an additional incentive mechanism that is relevant for real world negotiations.

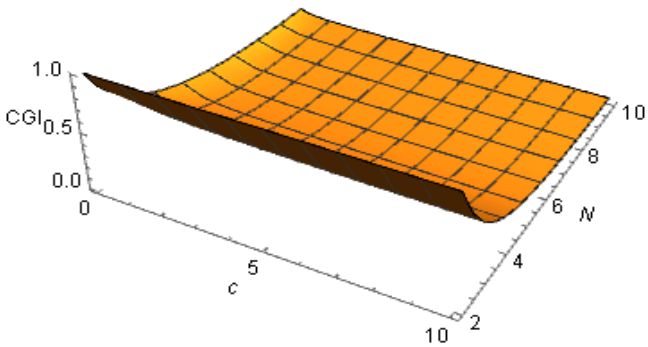


Figure 4.2: Efficiency of partial cooperation: numerical example 1

The relative efficiency of partial cooperation measured by the CGI is displayed for various simulated values of c and N with α=0.7.

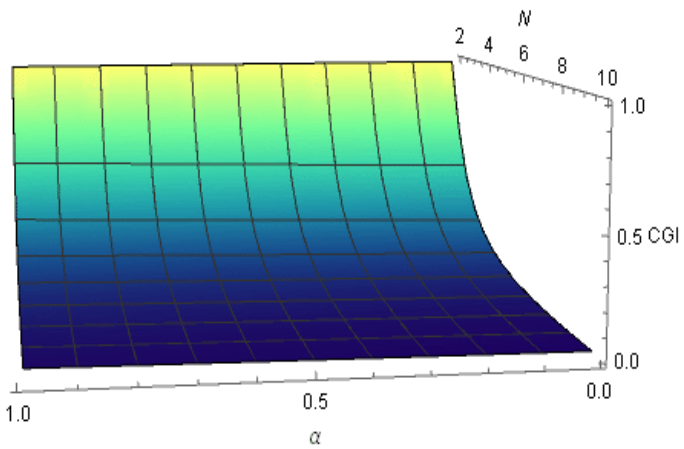


Figure 4.3: Efficiency of partial cooperation: numerical example 2

The relative efficiency of partial cooperation measured by the CGI is displayed for various simulated values for α and N with c=2.

4.4.3 The model of an international financial regulation agreement with exclusive club benefits

We extend the model by assuming that signatories additionally earn an exclusive club benefit. Until now, the public good characteristic of financial regulation is understood to be transmitted via the increased system-wide stability due to each individual country's efforts to foster a resilient financial system. In addition, it is plausible to assume additional benefits for signatories of the coalition, for instance in terms of more efficient information sharing, better market access or simply an increased mutual trust among countries. Goodhart (2011) notes for instance that information asymmetries are expected to decrease in the framework of regulatory cooperation by reducing costs for participating regulators. Furthermore, (financial) market access in most economies assumes the compliance with specific regulations. It might be required to be part of a regulatory cooperation in order to be allowed to operate on certain markets. Similarly, even though it might ultimately not be necessary to be part of a regulatory coalition, it may nevertheless be beneficial given that only signatories are able to set or influence the regulations in force. Being part of a regulatory cooperation, hence, can facilitate market access and increases a country's banks profitability. From another perspective, this interpretation is equivalent to the notion of decreasing costs of regulation for those countries participating in regulatory cooperation. Nordhaus (2015) shows that so-called climate clubs, which are able to induce internal club benefits or external sanctions on free riders are fruitful for combating climate change. In line with Nordhaus (2015), Hoel and Schneider (1997) show that assuming private non-environmental costs increases the stable size of environmental agreements. Therefore, it can be expected that the results of our model are positively influenced by additional club benefits even though it might be challenging to quantify this effect in practice.

We assume the club benefit of a signatory country to be given by:

$$E_s(n) = g(n-1). \tag{4.22}$$

For simplicity, this exclusive benefit is assumed to be a linear function of n. The only difference between this model version and the model presented in Chapter 4.4.2 is the benefit function of the signatories. The altered maximization problem of the signatories is given by:

$$\max_{q_s} \pi_s = \alpha\left(h\left(\frac{nq_s+(N-n)q_f}{N}\right) - \frac{1}{2}\left(\frac{nq_s+(N-n)q_f}{N}\right)^2\right) + (1-\alpha)\left(hq_s - \frac{1}{2}{q_s}^2\right) + g(n-1) - \frac{1}{2}c{q_s}^2 - \theta\left(q_s - \frac{nq_s+(N-n)q_f}{N}\right), \tag{4.23}$$

where $g \geq 0$. We notice that the club benefit does not depend on effort and thus effort levels in equilibrium are not affected. The net benefits of the signatories solely increase with the amount of club benefits. Thus, the difference between payoffs of signatories and free riders decreases which determines the stable size of the coalition. This implies that the incentive to sign a treaty increases with the existence of club benefits.

In order to analyze the stability of the grand coalition, we note from equation (4.20) that it is sufficient to use the internal stability condition $\pi_s(N) \geq \pi_f(N-1)$, because there are no countries outside the coalition $(n = N)$. We prove in the Technical Appendix C that the grand coalition is stable if the club benefits are sufficiently large. The notion of club benefits also allows to extend the findings in Dell'Ariccia and Marquez (2006) where the grand coalition could not be stable.

In order to measure the club benefits $E_s(N)$ required to stabilize the grand coalition we set them into relation to the amount of public benefits $H_s(N)$ which are given by:

$$H_s(N) = \alpha B_s(N), \quad \text{with} \quad B_s(N) = \frac{h^2(1+2c)}{2(1+c)^2}. \tag{4.24}$$

We use $\frac{E_s(N)}{H_s(N)} \equiv \gamma$ in order to have a more illustrative measure.

Proposition 5: For the model specification given by equation (4.4) including club benefits the self-enforcing IFRA has the following properties:

5.1 The existence of club benefits increases the stable size of a coalition and can even induce the grand coalition, if club benefits are sufficiently relevant.

5.2 The relation of club benefits to public benefits $\frac{E_s(N)}{H_s(N)}$ required in order to stabilize the grand coalition is strictly increasing in the competitiveness factor θ and strictly decreasing in the benefit factor h.

These are very important characteristics, since in practice we can observe regulatory cooperation including several countries. This result explains how club benefits could be an important driver to enlarge a stable coalition size over and above the pure internalization of externalities based on the public good characteristic of financial regulation. Proposition 5.2 gauges the relative size of the club good benefit $E_s(N)$ that is required to stabilize the grand coalition. We can see that stronger competition with other countries requires relatively larger club benefits to

induce the grand coalition while larger values of the benefit parameter h require a relatively smaller amount of club benefits.

We further illustrate how different magnitudes of club benefits influence the stable size of an agreement and how the existence of club benefits changes the influence of the other model parameters by using a numerical example. Figure 4.4 illustrates the stable size of a coalition depending on the public good parameter α and the club benefit using the same exogenous parameters values as in the previous simulation. First, we note that the coalition size n increases in the amount of club benefits.

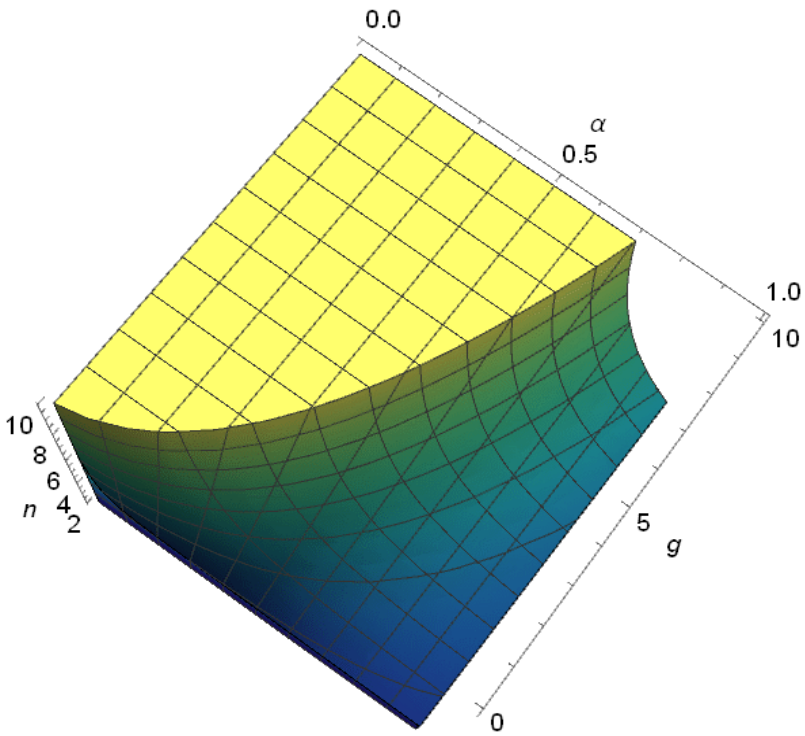


Figure 4.4: Stable coalition size: numerical example

The stable number of signatories n for various simulated values for α and g are displayed. The yellow area represents the grand coalition with n=N=10.

Second, we analyze if the existence of club benefits can alter the result that the coalition size is almost invariant in the public good parameter α. With $g > 0$ the size of the stable coalition decreases in α. However, for a given value of g the grand coalition ($n = N = 10$) is now stable for small values of α. This is due to the fact that the amount of club benefits is relatively high in relation to public benefits in such a situation. Club benefits are hence most efficient when cooperation is not that beneficial.

4.5 Model variant with linear benefit function

Naturally, our theoretical model and the derived results depend to an important extent on the chosen set-up. There is a trade-off between the necessary complexity of the modeling strategy and the generalization of results. So far, we have assumed a concave benefit function given that it is a realistic assumption that the marginal benefit from financial stability decreases. We advocate the view that it is desirable to assume a realistic set-up even though the complexity of the model might be considerably affected. However, this comes at certain cost because some interesting relations can only be formulated with the help of numerical illustrations. As an alternative, in this chapter we reduce the complexity of the model by assuming the linear benefit functions $B_i(\bar{q}) = b\bar{q}$ and $D_i(q_i) = bq_i$. On the one hand, this increases the level of abstractness and eliminates some important features of the previous setting. On the other hand, it allows deriving more general, analytical results by eliminating the necessity to simulate parameter constellations.

The altered net-benefits are given by:

$$\pi_i = \alpha b\bar{q} + (1-\alpha)bq_i - \frac{1}{2}c{q_i}^2 + g(n-1) - \theta(q_i - \bar{q}), \tag{4.25}$$

where $b > 0$. Linear benefits are special because they imply that reaction functions are orthogonal such that there is no strategic interaction. We can show that the main effects and results of the baseline setting are robust towards this alternative specification. We summarize the solution to the second stage in Lemma 1 with the proof given in the Technical Appendix C.

Lemma 1: Subgame perfect efforts and net benefits have the following properties:

1.1 The effort levels of free riders q_f^* are independent of n and the effort levels of signatories q_s^* are strictly increasing in n.

1.2 The net benefits of free riders π_f and signatories π_s are strictly increasing in n.

Then we solve stage one where the solution is characterized in Proposition 6 and the proof is given in the Technical Appendix C.

Proposition 6: For the self-enforcing agreement of the model specification given by equation (4.25), the following implications hold:

6.1 *If $g = 0$, two or three countries always cooperate.*

6.2 *If $g > 0$, depending on the amount of club benefits, coalitions from size three until the grand coalition exist. Incentives to cooperate strictly increase in c and N and strictly decrease in α and θ.*

6.3 *The relation of club benefits to public benefits $\frac{E_S(N)}{H_S(N)}$ required in order to stabilize the grand coalition is strictly increasing in the competitiveness factor θ and the number of countries N and strictly decreasing in the benefit factor b.*

6.4 *The CGI depends only on the number of countries N if $g = 0$ and depends on all $(N, n, b, \alpha, \theta)$ parameters if $g > 0$. The CGI converges to zero, if the number of countries N becomes large.*

Proposition 6.1 is in line with our previous results where we show that a coalition of $n = 2$ is stable for most parameter constellations. We see that this finding is not mainly driven by the assumed concave benefit functions, but that it is a robust finding for the assumed impure public good characteristic of financial stability. This is not very surprising since only extreme values come up with instable coalitions of size $n = 2$. Proposition 6.2 shows that the introduction of club benefits can once again lead to a stable coalition of any size including the grand coalition. The linear model specification can even induce further insights. Here, it can be shown that incentives to cooperate are directly and unanimously influenced by several variables when club benefits are taken into account. On the one hand, an increasing number of countries and increasing costs of regulation lead to a larger coalition size, because free-riding does not counteract increasing benefits from cooperation. On the other hand, the more strongly regulation is interpreted as a public good and the stronger competition with other countries can be assumed, the smaller a cooperative coalition will be.

The results of this chapter strengthen our view that large coalitions are indeed hard to establish but that at least small coalitions can be stable. The effect of a club benefit has similar implications as before and highlights a way to explain coalitions over and above three countries.

4.6 Discussion and relevance

The theoretical results derived in this work confirm widespread consensus that cooperation among financial supervisors is necessary to maintain and increase financial stability. The (impure) public good modelling of financial stability allows the assessment of the feasibility and efficiency of cooperative agreements. The results achieved so far indicate that regulatory cooperation in general is not very stable and not very large with a maximum number of signatories of two to three in the baseline settings. This result might only be surprising on first sight. However, when regarding cooperation rates among regulators on a global scale, it turns out that large regulatory agreements with strict and binding rules for each signatory are indeed hardly existent. The work of Beck et al. (2019) highlights this finding in a very detailed comprehensive empirical investigation. They find that, globally, only a small fraction of potential regulatory agreements is achieved and in case they are, the number of signatories is often rather small. Our work lays a rationale for this finding by providing theoretical evidence for the difficulty of reaching larger coalition sizes. In addition, our results indicate that given the obvious necessity to cooperate, smaller regulatory coalitions are indeed feasible and stable.

In addition, this work points a way to explain larger coalition sizes that are partly observable as stated in Beck et al. (2019), where clearly more than two or three countries cooperate. Club benefits might play a prominent role in order to stabilize larger coalitions. In fact, our modelling approach remains necessarily rather abstract and we cannot account for each specific influence that might be relevant in practice. Therefore, it is a straightforward approach to assume additional benefits that arise conditional on forming a coalition that exceed pure simple benefits from global supervisory efforts. This is a way to break down and handle considerably more complex relationships between countries by still incorporating the most important effects into our model.

The relevance of club benefits for the stability and size of a regulatory cooperation can also be interpreted in a normative sense. Given that regulatory cooperation is at the agenda of policy makers worldwide, club benefits might provide a promising approach to reach larger agreements. It could be the task of future research to investigate potential club benefits in more detail and to analyze the effective channels. This could allow policy makers to more directly address specific options that might be incorporated into negotiations regarding future supervisory agreements. For the time being, it is important to accentuate that additional club benefits are not only a way to explain a cooperation of a given size but also a potential remedy to overcome obstacles in the coalition formation process.

Finally, the choice of a home country regulation setting where each national supervisor overseas its national banks irrespective of the jurisdiction they operate in could be seen as a crucial assumption. As an alternative, host country regulation could be taken into account where each national supervisor overseas each bank in its own jurisdiction, irrespective of a bank being national or foreign. On the one hand, our model crucially depends on the home country assumption as this is the starting point for using the public good framework, such that results are not directly conceivable with host country regulation. On the other hand, however, we note that our model captures underlying patterns and mechanisms from the "micro-founded" bank-level model from Dell'Ariccia and Marquez (2006) and we could show that the results of our aggregate model mirror theirs. Therefore, we are optimistic that our results are also somewhat robust to a host country setting, following the arguments put forward in their work. They argue that even though in a host country setting foreign bank subsidiaries are regulated based on the regulation in the country of operation, there is still a strong link to the mother branch in the home country. The quality and level of capitalization of the bank subsidiary in the host country abroad is still strongly linked to the level and quality of capitalization of the mother company in the home country, which is determined by their own regulator at home. Hence, the general drivers of our model can be assumed to persist to some extent.

4.7 Conclusions

The cooperation of national regulators appears to be imperatively important given the increasing interconnections on global financial markets. In this work the cooperation of national financial regulators is studied in a game-theoretical framework. The provision of financial stability in terms of the so-called supervisory effort by regulators is interpreted as an impure public good, since aggregate efforts across several countries increase the overall financial stability while a certain benefit materializes only on local level. The public good character of financial regulation, however, leads to noticeable incentives to free-ride. A laxer national regulation translates into a competitive advantage of the domestic banking system. Similar issues have been largely studied in the literature on environmental economics, where cooperation incentives of countries are analyzed by methods from industrial organization. Due to the similarities of the underlying global public good structure, we apply solution methods from industrial organization. More concretely, we study the incentives of regulators to form self-enforcing international regulatory agreements that might prevent free-riding and increase the overall global financial stability.

Our approach enables us to determine the number of countries signing an agreement endogenously. We can show that partial cooperative solutions are feasible where, depending on the specific model set-up, at most two to three countries form a stable coalition. Our analysis also reveals that club benefits can raise the number of countries willing to cooperate significantly. This might point the way to policy makers aiming at increasing the size of self-enforcing regulatory agreements. Our results also confirm that cooperation is indeed desirable, since country-individual welfare maximization leads to an underprovision of financial regulation relative to a socially optimal situation where all countries jointly set regulatory standards. We show that full cooperation is hard to establish, and that potential free-riding behavior constitutes a real threat to joint efforts on a global level to maintain and improve financial stability. In general, those findings highlight the need to study how incentives of different national regulatory bodies might be aligned in order to increase the potential of common actions towards a more stable financial system.

Our model could be extended along various lines. It could be interesting to study how model parameters and results change when asymmetries among countries are considered. Moreover, it would be more realistic to assume multiple regulatory instruments in order to incorporate additional spillovers between them. A further extension could be to assume sequential effort decisions that allow studying potential first-mover (dis-) advantages. Additionally, the analysis of repeated interactions would be more realistic and could offer additional insights.

Economic research has a long tradition in studying the effects of public goods and free-riding behavior and it is advisable to resort to solution concepts from these strands of the literature, since they seem to be able to give valuable insights that can be used for the analysis of financial stability as well. In this exercise, we apply an impure public good modelling to the topic of financial regulation. The rather general model could be a valuable starting point for ample different topics that deal with cooperation on behalf of impure public goods.

5 Concluding remarks

The three essays brought together to form this dissertation have very distinct topics and methodologies, but all of them have the common unambiguous goal to study and analyze different facets of financial stability and to deepen the respective understanding that we possess. While they also show what we do not know (yet), several important conclusions and implications can still be drawn that are also shortly sketched in the following.

Essay 1 highlights the need for a careful application of systemic risk measures. We could show that different systemic risk measures tend to deliver diverging results regarding the systemic risk potential of financial institutions relative to other banks. Our analysis reveals the difficulties to denote specific bank-individual or macroeconomic characteristics that could be utilized for a better interpretation of systemic risk rankings. Our results eventually show that risk ranks are more strongly associated with bank and macroeconomic variables during low-volatility market phases, while they are seemingly unaffected during high-volatility phases. In addition, the specific channels that relate rank correlations to the set of variables remain often unclear. In fact, it should remind academics and practitioners not to rely too much on a single systemic risk metric but to take the results of several of them into account. More efforts are required to understand the underlying drivers of systemic risk and its respective measurement better.

The results of Essay 2 are also rather pessimistic with respect to an improved understanding of stock market volatility with the help of social media information. The various methods to estimate and forecast volatility cannot be significantly improved on average for a large sample of stocks when Twitter information is included. Still, there are many opportunities to widen the underlying analyses. Different ways to model Twitter information could be taken into account. Furthermore, a better understanding regarding the specific channels from Twitter to volatility needs a more detailed exploration. An important additional issue that is closely related concerns the few stocks where Twitter information seems to increase volatility forecast precision. Further analyses could shed more light on circumstances or specific characteristics that could induce Twitter information to become a more valuable source of information. Given the increasing availability of any kind of data due to the digitalization, this topic will certainly stay in the focus of academics and practitioners worldwide.

Finally, Essay 3 unfurls a stylized model of cooperation among financial supervisors. It can explain how difficult it is to establish and sustain joint regulatory policies due to contrarian incentives on the national level. It simultaneously highlights the need for more cooperation in order to accomplish and sustain a more stable financial system. Abstract in nature, it could

clearly be expanded by incorporating more concrete and more realistic aspects that do have an impact on regulatory cooperation. This could be an important approach when less general results should be obtained. Still, the chosen set-up permits to draw important conclusions that can serve as a valuable starting point for further analyses. The role of club benefits, for instance, is highly interesting given that it enables larger coalition sizes. The derivation of specific policy actions and recommendations is required to assess the nature, magnitude, and importance of these club benefits in practice.

References

Acharya, V. V. (2003): Is the international convergence of capital adequacy regulation desirable?, in: Journal of Finance, Vol. 58, No. 6, pp. 2745-2781.

Acharya, V.V., R. Engle, M. Richardson (2012): Capital shortfall: A new approach to ranking and regulation systemic risk, in: American Economic Review, Vol. 102, No. 3, pp. 59-64.

Acharya, V.V., L.H. Pedersen, T. Philippon, M. Richardson (2017): Measuring systemic risk, in: Review of Financial Studies, Vol. 30, No. 1, pp. 2-47.

Acharya, V.V., P. Schnabl, G. Suarez (2013): Securitization without risk transfer, in: Journal of Financial Economics, Vol. 107, No. 3, pp. 515-536.

Adrian, T., M.K. Brunnermeier (2016): CoVaR, in: American Economic Review, Vol. 106, No. 7, pp. 1705-41.

Agénor, P-R., L. Pereira da Silva (2018): Financial spillovers, spillbacks, and the scope for international macroprudential policy cooperation, in: BIS Working Paper No. 97.

Allen, F., S. Morris, H.Y. Shin (2006): Beauty contest and iterated expectations in asset markets, in: Review of Financial Studies, Vol. 19, No. 3, pp. 719-752.

Andersen, T.G., L. Benzoni, J. Lund (2002): An empirical investigation of continuous-time equity-return models, in: Journal of Finance, Vol. 73, No. 3, pp. 1239-1284.

Andersen, T.G., T. Bollerslev (1997) Intraday periodicity and volatility persistence in financial markets, in: Journal of Empirical Finance, Vol. 4, No. 2-3, pp. 115-158.

Andersen, T.G., T. Bollerslev (1998): Deutsche mark-dollar volatility: intraday activity patterns, macroeconomic announcements, and longer run dependencies, in: Journal of Finance Vol. 53, No. 1, pp. 219-265.

Andersen, T.G., Bollerslev, F.X. Diebold (2007): Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility, in: Review of Economics and Statistics, Vol. 89, No. 4, pp. 701-720.

Andersen, T.G., Bollerslev, F.X. Diebold, P. Labys (2003): Modeling and forecasting realized volatility, in: Econometrica, Vol. 71, No. 2, pp. 579-625.

Andersen, T.G., Bollerslev, P. Labys (2001): The distribution of realized exchange rate volatility, in: Journal of the American Statistical Association, Vol. 96, No. 453, pp. 42-55.

Anginer, D., A. Demirguc-Kunt, M. Zhu (2014): How does competition affect bank systemic risk?, in: Journal of Financial Intermediation, Vol. 23, No. 1, pp. 1-26.

Antweiler, W., M.Z. Frank (2004): Is all that talk just noise? The information content of internet stock message boards, in: Journal of Finance, Vol. 59, No. 3, pp. 1259-1294.

Bajgrowicz, P., O. Scaillet, A. Treccani (2016): Jumps in high-frequency data: Spurious detections, dynamics, and news, in: Management Science, Vol. 62, No. 8, pp. 2149-2455.

Barndorff-Nielsen, O.E., N. Shephard (2004): Power and bipower variation with stochastic volatility and jumps, in: Journal of Financial Econometrics, Vol.2, No. 1, pp. 1-37.

Barrett, S. (1994): Self-enforcing international environmental agreements, in: Oxford Economic Papers, Vol. 46, No. 0, pp. 878-894.

Barth, J.R., G. Caprio, R. Levine (2006): Rethinking banking regulation: Till angels govern, in: Cambridge University Press, New York.

Basel Committee on Banking Supervision (BCBS) (2013): Global systemically important banks: updated assessment methodology and the higher loss absorbency requirement, July, Bank for International Settlements.

Beck, T., O. Jonghe, K. Mulier (2018): Bank sectoral and (systemic) risk: Evidence from a worldwide sample of banks, Working Paper.

Beck, T., C. Silva-Buston, W. Wagner (2019): The economics of supranational bank supervision, in: CEPR Discussion Paper No. DP12764.

Behrendt, S., A Schmidt (2018): The Twitter myth revisited: Intraday investor sentiment, Twitter activity and individual-level stock return volatility, in: Journal of Banking and Finance, Vol. 96, pp. 355-367.

Beltratti, A., R.M. Stulz (2012): The credit crisis around the globe: Why did some banks perform better?, in: Journal of Financial Economics, Vol. 105, No. 1, pp. 1-17.

Benoit, S., G. Colletaz, C. Hurlin, C. Pérignon (2013): A theoretical and empirical comparison of systemic risk measures, Working Paper.

Benoit, S., J.-E. Colliard, C. Hurlin, C. Pérignon (2017): Where the risks lie: A survey on systemic risk, in: Review of Finance, Vol. 21, No. 1, pp. 109-152.

Benoit, S., C. Hurlin, C. Pérignon (2019): Pitfalls in systemic-risk scoring, in: Journal of Financial Intermediation, Vol. 38, pp. 19-44.

Benthaus, J., R. Beck (2015): It's more about the content than the users! The influence of social broadcasting on stock markets, in: ECIS 2015 Completed Research Papers, No. 17.

Bergstrom, T., L. Blume, H. Varian (1986): On the private provision of public goods, Journal of Public Economics, Vol. 29, No. 1, pp. 25-49.

Billio, M., M. Getmansky, A.W. Lo, L. Pelizzon (2012): Econometric measures of connectedness and systemic risk in the finance and insurance sectors, in: Journal of Financial Economics, Vol. 104, pp. 535-559.

Bollerslev, T., J. Cai, F. Song (2000): Intraday periodicity, long memory, and macroeconomic announcement effects in the US Treasury bond market, in: Journal of Empirical Finance, Vol. 7. No. 1, pp. 37-55.

Boudt, K., J. Cornelissen, S. Payseur (2013): Highfrequency: Toolkit for the analysis of highfrequency financialdata in R, working paper.

Bordino, I., S. Battiston, G. Caldarelli, M. Cristelli, A. Ukkonen, I. Weber (2012): Web search queries can predict stock market volumes, in: Plos One, Vol. 7, No. 7.

Bostandzic, D., M. Pelster, G.N.F. Weiß (2014): Systemic risk, bank capital, and deposit insurance around the world, Working Paper.

Brownlees, C., R. Engle (2011): Volatility, correlation and tails for systemic risk measurement, Working Paper.

Brownlees, C., R. Engle (2016): SRISK: A conditional capital shortfall index for systemic risk measurement, in: Review of Financial Studies, Vol. 30, No. 1, pp. 48-79.

Brunnermeier, M.K. (2009): Deciphering the liquidity and credit crunch 2007-2008, in: Journal of Economic Perspectives, Vol. 23, No.1, pp. 77-100.

Brunnermeier, M.K., G. Dong, D. Palia (2012): Banks‘ non-interest income and systemic Risk, Working Paper, Princeton University.

Brunnermeier, M.K., L.H. Pedersen (2009): Market liquidity and funding liquidity, in: Review of Financial Studies, Vol. 22, No. 6, pp. 2201-2238.

Buch, C. M., T. Krause, L. Tonzer (2017): Drivers of systemic risk: Do national and European perspectives differ?, in: Deutsche Bundesbank Discussion Paper, No. 09/2017.

Buck, F., E. Schliephake (2013): The regulator’s trade-off: bank supervision vs. minimum capital, in: Journal of Banking and Finance, Vol. 37, No. 11, pp. 4584-4598.

Cai, J., F. Eidam, A. Saunders, S. Steffen (2018): Syndication, interconnectedness, and systemic risk, in: Journal of Financial Stability, Vol. 34, pp. 105-120.

Cerchiello, P., P. Giudici, G. Nicola (2017): Twitter data models for bank risk contagion, in: Neurocomputing, Vol. 264, pp. 50-56.

Chiriac, R., V. Voev (2011): Modelling and forecasting multivariate realized volatility, in: Journal of Applied Econometrics, Vol. 26, No. 6, pp. 922-947.

Cihák, M., A. Demirgüc-Kunt, M. Martinez Peria, A. Mohseni-Cheraghlou (2013): Bank regulation and supervision in the context of the global crisis, in: Journal of Financial Stability, Vol. 9, No. 4, pp. 733-746.

Cont, R., (2001): Empirical properties of asset returns: stylized facts and statistical issues, in: Quantitative Finance, Vol. 1, No. 2, pp. 223-236.

Corsi, F. (2009): A simple approximate long-memory model of realized volatility, in: Journal of Financial Econometrics, Vol. 7, No. 2, pp. 174-196.

Da, Z., J. Engelberg, P. Gao (2015): The sum of all fears: investor sentiment and asset prices, in: Review of Financial Studies, Vol. 28, No. 1, pp. 1-32.

Danielsson, J., K. James, M. Valencuela, I. Zer (2016a): Model risk of risk models, in: Journal of Financial Stability, Vol. 23, pp. 79-91.

Danielsson, J., K. James, M. Valencuela, I. Zer (2016b): Can we prove a bank guilty of creating systemic risk? A minority report, in: Journal of Credit, Money and Banking, Vol. 48, No. 4, pp. 795-812.

D'Aspremont, C.A., J. Jacquemin, J. Gabezeweiz, J.A. Weymark (1983): On the stability of collusive price leadership, in: Canadian Journal of Economics, Vol. 16, No. 1, pp. 17-25.

Degryse, H., S. Ongena (2005): Distance, lending relationships and competition, in: Journal of Finance, Vol. 60, No. 1, pp. 231-266.

De Jonghe, O., H. Dewachter, S. Ongena (2020): Bank capital (requirements) and credit supply: evidence from Pillar 2 decisions, in: Journal of Corporate Finance, Vol. 60.

Dell'Ariccia, G., R. Marquez (2001): Competition among regulators, in: IMF Working Paper WP/01/73.

Dell'Ariccia, G., R. Marquez (2006): Competition among regulators and credit market integration, in: Journal of Financial Economics, Vol. 79, No. 2, pp. 401-430.

De Long, J.B., A. Shleifer, L. Summers, R.J. Waldmann (1990: Noise trader risk in financial markets, in: Journal of Political Economy, Vol. 98, No. 4, pp. 703-738.

Demirgüç-Kunt, A., E. Detragiache (2002): Does deposit insurance increase banking system stability? An empirical investigation, in: Journal of Monetary Economics, Vol. 49, No. 7, pp. 1373-1406.

Diebold, F.X., R.S. Mariano (1995): 'Comparing predictive accuracy', in: Journal of Business and Economic Statistics, Vol. 13, No.3, pp. 134–44.

Diebold, F.X. (2015): Comparing predictive accuracy, twenty years later: A personal perspective on the use and abuse of Diebold–Mariano tests, in: Journal of Business and Economic Statistics, Vol. 33, No.1, pp. 1-9.

Dimpfl, T., S. Jank (2016): Can internet search queries help to predict stock market volatility?, in: European Financial Management, Vol. 22, No. 2, pp. 171-192.

Döring, B., C. Wewel, T. Hartmann-Wendels (2016): Systemic risk measures and their viability for banking supervision, Working Paper, Department of Bank Management, University of Cologne.

Eldridge, D.S., H.H. Ryoo, A. Wieneke (2015): Bank capital regulation with asymmetric countries, in: Economic Record, Vol. 91, No. 292, pp. 79-90.

Ellul, A., C.W. Holden, P. Jain, R. Jennings (2007): Order dynamics: Recent evidence from the NYSE, in: Journal of Empirical Finance, Vol. 14, No. 5, pp. 636-661.

Engle, R. (2002): Dynamic conditional correlation: A simple class of multivariate gener-

alized autoregressive conditional heteroskedasticity models, in: Journal of Business and Economic Statistics, Vol. 20, No. 3, pp. 339-350.

Engle, R.F., E. Jondeau, M. Rockinger (2015): Systemic risk in Europe, in: Review of Finance, Vol. 19, pp. 145-190.

Engle, R.F., M. E. Sokalska (2012): Forecasting intraday volatility in the US equity market. Multiplicative component GARCH, in: Journal of Financial Econometrics, Vol. 10, No. 1, pp. 54-83.

Eyckmans, J., M. Finus (2006): Coalition formation in a global warming game: how the design of protocols affects the success of environmental treaty-making, in: Natural Resource Modeling, Vol. 19, No. 3, pp. 323-358.

Fama, E. F. (1970): Efficient capital markets: A review of theory and empirical work, in: Journal of Finance, Vol. 25, No. 2, pp. 383-417.

Fan, R., O. Talavera, V. Tran (2019): Social media bots and stock markets, forthcoming in: European Financial Management.

Finus, M., D.T. Rübbelke (2013): Public good provision and ancillary benefits: The case of climate agreements, in: Environmental and Resource Economics, Vol. 56, No. 2, pp. 211-226.

Gallant, A.R. (1981): On the bias in flexible functional forms and an essentially unbiased form, in: Journal of Econometrics, Vol. 15, No. 2, pp. 211-245.

Gandhi, P., H. Lustig (2015): Size anomalies in US bank stock returns, in: Journal of Finance, Vol. 70, No. 2, pp. 733-768.

Gaspar, V., G. Schinasi (2010): Financial stability and policy cooperation, in: Bank of Portugal Occasional Paper 1/2010.

Gehrig, T. (2014), SOAS Centenary Lecture on International Banking Supervision.

Giglio, S., B.T. Kelly, S. Pruitt (2016): Systemic risk and the macroeconomy: An empirical evaluation, in: Journal of Financial Economics, Vol. 119, No. 3, pp. 457-471.

Girardi, G., A.T. Ergün (2013): Systemic risk measurement: Multivariate GARCH estimation of CoVaR, in: Journal of Banking and Finance, Vol. 37, No. 8, pp. 3169-3180.

Goodhart, C. (2011): The Basel Committee on Banking Supervision: a history of the early years, 1974-1997, Cambridge University Press, Cambridge.

Gosh, A., J. Ostry, M. Qureshi (2017): Managing the tide: how do emerging markets respond to capital flows?, in: IMF Working Paper No. 17/69.

Gravelle, T., F. Li (2013): Measuring systemic importance of financial institutions: An extreme value theory approach, in: Journal of Banking and Finance, Vol. 37, pp. 2196-2209.

Grundke, P. (2019): Banking network models and the ranking consistency of systemic risk

measures: A simulation-based analysis, in: Review of Quantitative Finance and Accounting, Vol. 52, No. 4, pp. 953-990.

Grundke, P., M. Tuchscherer (2019): Global systemic risk measures and their forecasting power for systemic events, in: European Journal of Finance, Vol. 25, No. 3, pp. 205-233.

Hansen, P.R., A Lunde (2006): Realized variance and market microstructure noise, in: Journal of Business & Economic Statistics, Vol. 24, No. 2, pp. 127-161.

Hardy, D.C., M.J. Nieto (2011): Cross-border coordination of prudential supervision and deposit guarantees, in: Journal of Financial Stability, Vol.7, No. 3, pp. 155-164.

Hauswald, R., R. Marquez (2006), Competition and strategic information acquisition in credit markets, in: Review of Financial Studies, Vol. 19, No. 3, pp. 967-1000.

Hautsch, N. (2012): Econometrics of financial high-frequency data, Springer, Berlin Heidelberg.

Hautsch, N., M. Podolskij (2013): Preaveraging-based estimation of quadratic variation in the presence of noise and jumps: theory, implementation, and empirical evidence, in: Journal of Business and Economic Statistics, Vol. 31, No. 2, pp. 165-183.

Hoel, M. and Schneider, K. (1997), Incentives to Participate in an International Environmental Agreement, in: Environmental and Resource Economics, Vol. 9, pp. 153-170.

Holthausen, C., T. Rønde (2005): Cooperation in international banking supervision, in: CEPR Discussion Paper Series No. 4990.

Homar, T., H. Kick, C. Sallelo (2016): Making sense of the EU wide stress test: a comparison with the SRISK approach, ECB Working Paper No 1920.

Hovakimian, A., E. Kane, L. Laeven (2012): Variation in systemic risk at US banks during 1974-2010, NBER Technical Report.

Huang, X., G. Tauchen (2005): The relative contribution of jumps to total price variance, in: Journal of Financial Econometrics, Vol. 3, No.4, pp. 456-499.

Irresberger, F., C. Bierth, G.N.F. Weiß (2017): Size is everything: Explaining SIFI designations, in: Review of Financial Economics, Vol. 32, pp. 7-19.

Kahneman, D., A. Tversky (1979): Prospect theory: An analysis of decision under risk, in: Econometrica, Vol. 47, No. 2, pp. 263-291.

Kara, G. (2016): Systemic risk, international regulation and the limits of coordination, in: Journal of International Economics, Vol. 99, pp. 192-222.

Karppi, T. and K. Crawford (2016): Social Media, Financial Algorithms and the Hack Crash, in: Theory, Culture and Society, Vol. 33, No. 1, pp. 73-92.

Kaul, I., I. Grunberg, M. Stern (1999):, Global public goods: International cooperation in the 21st Century, Oxford University Press, New York.

Kendall, M., J. Gibbons (1990): Rank correlation methods, 5th edition, Edward Arnold Publishing, London.

Korinek, A. (2017): Currency wars or efficient spillovers? A general theory of international policy cooperation, in: IMF Working Paper No. 17/25.

Laeven, L., L. Ratnovski, H. Tong (2016): Bank size, capital, and systemic risk: Some international evidence, in: Journal of Banking and Finance, Vol. 69, pp. 25-34.

Li, X., D. Tripe, C. Malone (2017): Measuring bank risk: An exploration of z-score, Working Paper, Massey University, New Zealand.

Löffler, G., P. Raupach (2018): Pitfalls in the use of systemic risk measures, in: Journal of Financial and Quantitative Analysis, Vol. 53, N0. 1, pp. 269-298.

López-Espinosa, G., A. Moreno, A. Rubia, L. Valderrama (2012): Short-term wholesale funding and systemic risk: A global CoVaR approach, in: Journal of Banking and Finance, Vol. 36, No. 12, pp. 3150-3162.

López-Espinosa, G., A. Moreno, A. Rubia, L. Valderrama (2015): Systemic risk and asymmetric responses in the financial industry, in: Journal of Banking and Finance, Vol. 58, pp. 471-485.

Mao, H., S. Counts, J. Bollen (2011): Predicting financial markets: Comparing survey, news, Twitter and search engine data, Working Paper.

Mao, H., S. Counts, J. Bollen (2015): Quantifying the effects of online bullishness on international financial markets, in: ECB Statistics Paper Series, No. 9.

Mühlnickel, J., G.N.F. Weiß (2015): Consolidation and systemic risk in the international insurance industry, in: Journal of Financial Stability, Vol. 18, pp. 187-202.

Nordhaus, W. (2015): Climate clubs: overcoming free-riding in international climate policy, in: American Economic Review, Vol. 105, No. 4, pp. 1339-1370.

Nucera, F., B. Schwaab, S. J. Koopmann, A. Lucas (2016): The information in systemic risk rankings, in: Journal of Empirical Finance, Vol. 38, pp. 461-175.

Olson, M., R. Zeckhauser (1966): An economic theory of alliances, in: Review of Economics and Statistics, Vol. 48, No. 3, pp. 266-279.

Park, S., Y.-H. Kim (2018): International policy coordination for financial regime stability under cross-border externalities, in: Journal of Banking and Finance, Vol. 97, pp. 177-188.

Patton, A.J. (2011): Volatility forecast comparison using imperfect volatility proxies, in: Journal of Econometrics, Vol. 160, No. 1, pp. 246-256.

Poon, S.H., Granger, C.W. (2003): Forecasting volatility in financial markets: a review, in: Journal of Economic Literature, Vol. 41, No. 2, pp. 478-539.

Prokopczuk, M., L. Symeonidis, C. Wese Simen (2016): Do jumps matter for volatility

forecasting? Evidence from energy markets, in: Journal of Future Markets, Vol. 36, No. 8, pp. 759-792.

Ranco, G., D. Aleksovski, G. Caldarelli, M. Grcar, I. Mozetic (2015): The effects of Twitter sentiment on stock price returns, in: Plos One, Vol. 10, No. 9.

Renault, T. (2017): Intraday online investor sentiment and return patterns in the US stock market, in: Journal of Banking and Finance, Vol. 84, pp. 25-40.

Revesz, Rachael (2016): Donald Trump wipes off $1bn from Boeing's stock market value with single tweet, in: The Independent, Online article: https://www.independent.co.uk/news/business/donald-trump-targets-boeing-shares-fall-twitter-air-force-one-planes-dennis-muilenburg-a7458946.html, retrieved on 29 March 2020.

Scaillet, O. (2005): Nonparametric estimation of conditional expected shortfall, in: Insurance and Risk Management Journal, Vol. 74, No. 1, pp. 639-660.

Schoenmaker, D. (2011): The financial trilemma, in: Economics Letters, Vol. 111, No. 1, pp. 57-59.

Sinn, H.-W. (2003): Risk taking, limited liability, and the competition of bank regulators, FinanzArchiv: Public Finance Analysis, Vol. 59, No. 3, pp. 305-329.

Souza, T., T. Aste (2016): A nonlinear impact: evidences of causal effects of social media on market prices, Working paper.

Sprenger, T.O., P.G. Sandner, A. Tumasjan, I. Welpe (2013): Tweets and trades: The information content of stock microblogs, in: European Financial Management, Vol. 20. No.5, pp. 926-957.

Sprenger, T.O., P.G. Sandner, A. Tumasjan, I. Welpe (2014): News or noise? Using Twitter to identify and understand company-specific news flow, in: Journal of Business Finance and Accounting, Vol. 41, No. 7, pp. 791-830.

Tetlock, P.C. (2007): Giving content to investor sentiment: The role of media in the stock market, in: Journal of Finance, Vol. 62, No. 3, pp. 1139-1168.

Uhde, A., U. Heimeshoff (2009): Consolidation in banking and financial stability in Europe: Empirical evidence, in: Journal of Banking and Finance, Vol. 33, pp. 1299-1311.

Uluc, A., T. Wieladek (2017): Capital requirements, risk shifting and the mortgage market, in: European Central Bank Working Paper Series No 2061.

Vallascas, F. J. Hagendorff (2011): The impact of European bank mergers on bidder default risk, in: Journal of Banking and Finance, Vol. 35, No. 4, pp. 902-915.

Van Hoose, D. (2016): Should financial regulators engage in international policy coordination?, International Economics and Economic Policy, Vol. 13, No. 2, pp. 319-338.

Weiß, G.N.F., S. Neumann, D. Bostandzic (2014a): Systemic risk and bank consolidation: International Evidence, in: Journal of Banking and Finance, Vol. 40, No. 3, pp. 165-181.

Weiß, G.N.F., D. Bostandzic, S. Neumann (2014b): What factors drive systemic risk during international financial crisis?, in: Journal of Banking and Finance, Vol. 41, No. 4, pp. 78-96.

White, L. J. (1994): On the international harmonization of bank regulation, in: Oxford Review of Economic Policy, Vol. 10, No. 4, pp. 94-105.

Wysocki, P.D. (2000): Private information earning announcements and trading volume or stock chat on the Internet: A public debate about private information, Working paper, Ann Arbor University of Michigan Business School.

Zhang, X., H. Fuehres, P. Gloor (2011): Predicting stock market indicators through Twitter “I hope it is not as bad as I fear", in: Procedia- Social and Behavioral Sciences, Vol. 26, pp. 55-62.

Zheludev, I., R. Smith, T. Aste (2014): When can social media lead financial markets?, in: Nature - Scientific Reports Vol. 4, No. 4213, pp. 1-12.

Zhou, C., N. Tarashev (2013): Looking at the tail: price-based measures of systemic importance, in: BIS Quarterly Review, June.

Appendix A: (Essay 1)

Table A.1:List of banks

Aareal Bank	Goldman Sachs
Allied Irish Banks	HSBC
Alpha Bank	Hudson City
American Express	Huntington Bancshares
Annaly Capital Management	ING
Banca Monte dei Paschi di Siena	Intesa Sanpaolo
Banca Popolare di Milano	JP Morgan Chase
Banco BPI	Jyske Bank
Banco Comercial Portugues	KBC Group
Banco de Sabadell	KeyCorp
Banco Popular Espanol	Lloyds Banking Group
Banco Santander	Mediobanca
Bank of America	Morgan Stanley
Bank of Ireland	M&T Bank
Bank of Montreal	National Bank of Greece
Bank of New York Mellon	Natixis
Bank of Nova Scotia	Nordea
Bank of Piraeus	Northern Trust
Bankinter	Permanent TSB
Barclays	PNC Financial Services
BB&T	Raiffeisen Bank
BBVA	Regions Financial
BNP	Royal Bank of Canada
Caixabank	Royal Bank of Scotland
Canadian Imperial Bank of Commerce	SEB
Capital One Financial	SNS Reaal
Charles Schwab	Societe Generale
Citigroup	Standard Chartered
Comerica	State Street
Commerzbank	Storebrand
Credit Agricole	SunTrust
Credit Suisse	Svenska Handelsbanken
Danske Bank	Swedbank
Deutsche Bank	Sydbank
Dexia	UBS
Discover Financial Services	UniCredit
DNB	Unione di Banche Italiane
Erste Group	US Bancorp
Eurobank Ergasias	Wells Fargo
Fifth Third Bank	Zions Bancorporation

This table contains all banks of the sample.

Table A.2: Literature overview: Variables that can be associated with the level of systemic risk measures

Variable	SRM	Direction	Statistical significance	Source
Total assets	MES	Positive	Insignificant	Weiß et al. (2014a)
		Positive	Insignificant	Bostandzic et al. (2014)
		Mixed	Insignificant	Weiß et al. (2014b)
	SRISK	Positive	Significant	Bostandzic et al. (2014)
		Positive	Insignificant	Homar et al. (2016)
		Positive	Significant	Buch et al. (2017)
		Positive	Significant	Laeven et al. (2016)
	ΔCoVaR	Positive	Insignificant	López-Espinosa et al. (2012)
		Positive	Significant	Girardi and Ergün (2013)
		Positive	Insignificant	Bostandzic et al. (2014)
		Positive	Significant	Adrian and Brunnermeier (2016)
		Positive	Significant	Laeven et al. (2016)
Loan loss provision	MES	Neutral	Weakly significant	Weiß et al. (2014a)
		Positive	Significant	Bostandzic et al. (2014)
		Neutral	Insignificant	Weiß et al. (2014b)
	SRISK	Positive	Significant	Bostandzic et al. (2014)
	ΔCoVaR	Mixed	Insignificant	Bostandzic et al. (2014)
Non-performing loans to total loans	MES	Positive	Insignificant	Döring et al. (2016)
	SRISK	Positive	Insignificant	Döring et al. (2016)
		Positive	Insignificant	Buch et al. (2017)
Market-to-book ratio	MES	Negative	insignificant	Weiß et al. (2014a)
		Positive	Insignificant	Bostandzic et al. (2014)
		Mixed	Partly insignificant	Weiß et al. (2014b)
		Negative	Insignificant	Döring et al. (2016)
	SRISK	Mixed	Insignificant	Bostandzic et al. (2014)
		Negative	Significant	Döring et al. (2016)
	ΔCoVaR	Neutral	Insignificant	López-Espinosa et al. (2012)
		Negative	Insignificant	Bostandzic et al. (2014)
		Negative	Significant	Döring et al. (2016)
Leverage	MES	Positive	Partly significant	Bostandzic et al. (2014)
		Neutral	Insignificant	Weiß et al. (2014b)
		Negative	Insignificant	Döring et al. (2016)
	SRISK	Positive	Insignificant	Bostandzic et al. (2014)
		Negative	Significant	Döring et al. (2016)
		Positive	Insignificant	Homar et al. (2016)

Table A.2 (cont.)

	ΔCoVaR	Positive/neutral Neutral Negative	Partly significant (during market downturns) Insignificant Insignificant	Girardi and Ergün (2013) Bostandzic et al. (2014) Döring et al. (2016)
ROA	MES	Positive	Insignificant	Weiß et al. (2014a)
	SRISK	Mixed Positive	Insignificant Weakly significant	Homar et al. (2016) Buch et al. (2017)
Long-term funding to total funding	MES	Positive	Insignificant	Bostandzic et al. (2014)
	SRISK	Mixed	Insignificant	Bostandzic et al. (2014)
	ΔCoVaR	Negative	Insignificant	Bostandzic et al. (2014)
Non-interest to total interest income	MES	Mixed Negative Neutral	Insignificant Insignificant Insignificant	Weiß et al. (2014a) Bostandzic et al. (2014) Weiß et al. (2014b)
	SRISK	Mixed Negative	Insignificant Insignificant	Bostandzic et al. (2014) Buch et al. (2017)
	ΔCoVaR	Positive	Weakly significant	Bostandzic et al. (2014)
Loans to deposits	MES	Positive	Weakly significant	Döring et al. (2016)
	SRISK	Positive	Significant	Döring et al. (2016)
	ΔCoVaR	Positive	Significant	Döring et al. (2016)
Tier1 ratio	MES	Negative	Insignificant	Bostandzic et al. (2014)
	SRISK	Positive Positive	Insignificant Significant	Bostandzic et al. (2014) Laeven et al. (2016)
	ΔCoVaR	Positive Positive	Significant Insignificant	Bostandzic et al. (2014) Laeven et al. (2016)
Interconnectedness	MES	Positive	Weakly significant	Bostandzic et al. (2014)
	SRISK	Positive	Weakly Significant	Bostandzic et al. (2014)
	ΔCoVaR	Negative	Significant	Bostandzic et al. (2014)
Unemployment	MES	Negative	Insignificant	Döring et al. (2016)
	SRISK	Positive	Insignificant	Döring et al. (2016)
	ΔCoVaR	Negative	Significant	Döring et al. (2016)
Inflation	MES	Neutral Positive	Weakly significant Significant	Weiß et al. (2014b) Döring et al. (2016)
	SRISK	Positive Negative	Significant Insignificant	Döring et al. (2016) Buch et al. (2017)
	ΔCoVaR	Positive	Weakly significant	Döring et al. (2016)

Table A.2 (cont.)

GDP growth	MES	Positive	Insignificant	Weiß et al. (2014a)
		Negative	Weakly significant	Bostandzic et al. (2014)
		Mixed	Significant	Weiß et al. (2014b)
		Positive	Insignificant	Döring et al. (2016)
	SRISK	Negative	Weakly significant	Bostandzic et al. (2014)
		Positive	Insignificant	Döring et al. (2016)
		Negative	Insignificant	Buch et al. (2017)
	ΔCoVaR	Neutral	Insignificant	Bostandzic et al. (2014)
		Neutral	Insignificant	Döring et al. (2016)
HHI	MES	Mixed	Significant	Weiß et al. (2014b)
		Positive	Significant	Bostandzic et al. (2014)
	SRISK	Positive	Significant	Bostandzic et al. (2014)
	ΔCoVaR	Positive	Weakly significant	Bostandzic et al. (2014)
Stock index return	ΔCoVaR	Mixed	Significant	Adrian and Brunnermeier (2016)
Stock index volatility	ΔCoVaR	Positive	Significant	Adrian and Brunnermeier (2016)

This table shows an overview of variables whose association with the levels of the different systemic risk measures have been tested in the literature. SRM indicates the specific systemic risk measure, Direction indicates if the SRM is positively or negatively associated (or not all) and statistical significance states the statistical significance of the relation.

Table A.3: Variable description

Variable name	Description	Source
Bank variables		
Assets	Natural logarithm of total assets (in thousands of US-Dollar)	Bankscope
Reserves-to-loans	Reserves against possible losses on impaired or non-performing loans/total loans (in thousands of US-Dollar)	Bankscope, Thomson Reuters Eikon (market capitalization)
Market-to-book	Market value equity (measured as market capitalization)/book value total equity (in thousands of US-Dollar)	Bankscope, Thomson Reuters Eikon (market capitalization)
Leverage	Ratio of total debt and total assets	Bankscope
ROA	Return on assets	Bankscope
Long-term funding	Long-term funding/total funding (in thousands of US-Dollar)	Bankscope
Non-interest income	Non-interest income/total interest income (in thousands of US-Dollar)	Bankscope
Loans-to-deposits	Total loans/total (customer) deposits (in thousands of US-Dollar)	Bankscope
Tier 1 ratio	Tier 1 capital ratio	Bankscope
Z-score	((Total equity/total assets)+ROA)/Std.Dev(ROA)	Bankscope
Interconnectedness	The ratio of the sum of in- and out-connections of standardized bank equity returns based on a Granger causality test as developed in Billio et al. (2012) to the total amount of potential connections.	Thomson Reuters Eikon, own calculations
Total liabilities	Total liabilities (in thousands of US-Dollar)	Bankscope
Equity (stock) prices	Daily closing prices of the respective banks	Thomson Reuters Eikon

Table A.3 (cont.)

MSCI world index	Daily closing prices of MSCI world index	Thomson Reuters Eikon
Macro variables		
Unemployment	Unemployment as a fraction of total labor force	WDI Database
Inflation	Change in consumer price index (in %)/100	WDI Database
GDP growth	Nominal GDP growth rate at market prices in local currency (in %)/100	WDI database
Herfindahl-Hirschman-Index (HHI)	Sum of squared market shares based on total assets of a country's domestic and foreign banks	Thomson Reuters Eikon, ECB, national central banks, own calculations
Long-term interest rate	10-years government bond yield (in %)	IMF, OECD
Stock index return	Quarterly averages of daily log-returns of a country's major stock index (in %)	Thomson Reuters Eikon
Stock index volatility	Quarterly averages of daily annualized 22-days log-return volatility of a country's major stock index.	Thomson Reuters Eikon
VIX	Implied volatility index	Thomson Reuters Eikon

This table contains all the variables used as exogenous regressors in the panel analysis as well as for the computation of the systemic risk measures. All bank variables are – if not indicated differently – at a quarterly frequency. Unemployment, Inflation and GDP growth are at annual frequency and transformed to quarterly values via cubic splines. The remaining macroeconomic variables are at daily frequency and transformed to quarterly averages. The sample range covers the period 2009 Quarter 3 to 2016 Quarter 3.

Table A.4: Summary statistics - Baseline regression

Variable	Mean	Median	Std. Dev.	Min	25%-Quantile	75%-Quantile	Max	Significance
Exposure systemic risk measures								
Corr MES-SRISK	0.3910	0.4607	0.3725	-0.8982	0.1587	0.6897	0.9738	68% (76%)
Corr MES-MES-EQ	0.6840	0.7576	0.2508	-0.5617	0.5791	0.8684	0.9847	91% (93%)
Corr MES-ExpΔCoVaR	0.9508	0.9686	0.0595	0.3572	0.9434	0.9825	1.0000	100%
Corr MES-ExpΔCoVaR-EQ	0.6850	0.7604	0.2526	-0.5487	0.5880	0.8665	0.9877	91% (93%)
Corr SRISK-MES-EQ	0.2908	0.3454	0.3860	-0.8833	0.0219	0.5970	0.9674	60% (70%)
Corr SRISK-ExpΔCoVaR	0.3859	0.4522	0.3738	-0.8894	0.1610	0.6874	0.9697	67% (76%)
Corr SRISK-ExpΔCoVaR-EQ	0.2915	0.3474	0.3857	-0.9072	0.0329	0.5935	0.9635	61% (70%)
Corr MES-EQ-ExpΔCoVaR	0.6757	0.7513	0.2540	-0.5649	0.5660	0.8622	0.9859	90% (93%9
Corr MES-EQ-ExpΔCoVaR-EQ	0.9042	0.9314	0.0936	-0.7016	0.8780	0.9619	1.0000	100%
Corr ExpΔCoVaR –ExpΔCoVaR-EQ	0.6819	0.7595	0.2529	-0.5264	0.5790	0.8634	0.9890	91% (93%)
Explanatory variables								
Assets	19.5199	19.4523	1.2366	17.1232	18.4493	20.5458	21.7223	
Reserves-to-loans	0.0317	0.0222	0.0333	0.0020	0.0109	0.0392	0.1891	
Market-to-book	0.9878	0.8551	0.6641	0.0083	0.5694	1.2107	4.0697	
Leverage	0.93048	0.9352	0.0305	0.8579	0.9087	0.9537	0.9940	
ROA	0.00150	0.0017	0.0039	-0.0175	0.0005	0.0030	0.0126	
Long-term funding	0.1782	0.1423	0.1354	0.0066	0.0823	0.2520	0.6147	

Table A.4 (cont.)

Non-interest income	0.7383	0.4467	1.1041	-1.2468	0.2319	0.7111	6.0433
Loans-to-deposits	1.2684	1.1266	0.6605	0.10650	0.8649	1.5313	4.7115
Tier1 ratio	0.1256	0.1210	0.0322	-0.0690	0.1080	0.1390	0.3587
Z-score	52.7103	43.9736	40.5247	0.0967	21.3351	71.6046	190.0186
Interconnectedness	0.6761	0.6835	0.1003	0.3987	0.6139	0.7468	0.8671
Unemployment	0.0945	0.0801	0.0524	0.0307	0.0670	0.0984	0.2684
Inflation	0.0133	0.0147	0.0122	-0.0197	0.0034	0.0223	0.0421
GDP growth	0.0039	0.0081	.0237	-0.0731	-0.0089	0.0203	0.0587
HHI	0.0835	0.0527	0.0552	0.0273	.05116	0.1068	0.2254
Long-term interest rate	3.1103	2.6233	2.2977	0.1400	1.9067	3.5833	15.4967
Stock index return	0.0288	0.0378	0.1641	-0.4582	-0.0531	0.1292	0.4380
Stock index volatility	0.2150	0.1897	0.0981	0.0842	0.1395	0.2666	0.5087

For a detailed description of the employed variables, see Table 3. All of the rank correlation pairs are computed based on one quarter of daily ranks for each systemic risk ranking and each individual bank for a sample. The column “Significance” indicates the share of correlations that are significantly different from zero at the 99% (95%) level. The sample of banks contains 80 large international banks. Data range is from quarter 2, 2009 to quarter 3, 2016. The explanatory variables are winsorized at the 1% and 99% level.

Table A.5: Regression results for more volatile market phases

	MES-MES-EQ	MES-SRISK	MES-ExpΔCo-VaR	MES-ExpΔCo-VaR-EQ	MES-EQ-SRISK	MES-EQ - ExpΔCo-VaR	MES-EQ-ExpΔCo-VaR-EQ	SRISK-ExpΔCo-VaR	SRISK-ExpΔCo-VaR EQ	ExpΔCo-VaR-ExpΔCo-VaR-EQ
Assets	0.123**	-0.188*	-0.0031	0.0772	-0.0602	0.125**	0.0269	-0.167*	-0.0940	0.0796
	(2.05)	(-1.91)	(-0.35)	(1.38)	(-0.60)	(2.15)	(1.52)	(-1.74)	(-0.93)	(1.45)
Reserves-to-loans	-0.791	-0.774	-0.0236	-1.057	-0.144	-1.012	0.205	-0.997	-0.0505	-1.276
	(-0.86)	(-0.73)	(-0.24)	(-1.19)	(-0.10)	(-1.02)	(0.96)	(-0.91)	(-0.03)	(-1.35)
Market-to-book	0.0329	0.0980	-0.0061	0.0338	0.103	0.0356	0.0101	0.0918	0.0759	0.0368
	(0.97)	(1.53)	(-0.72)	(1.07)	(1.44)	(1.05)	(0.87)	(1.50)	(1.10)	(1.22)
Leverage	-0.784	-0.0564	-0.0389	-0.668	-0.130	-0.734	0.265	-0.0593	-0.0365	-0.625
	(-0.96)	(-0.07)	(-0.47)	(-0.78)	(-0.12)	(-0.86)	(1.21)	(-0.08)	(-0.03)	(-0.72)
ROA	-0.0169	-5.097	-0.0711	0.943	-1.011	-0.133	1.048	-5.172	-0.712	0.540
	(-0.01)	(-0.80)	(-0.15)	(0.35)	(-0.16)	(-0.06)	(1.67)	(-0.84)	(-0.11)	(0.21)
Long-term funding	0.168	0.392*	-0.0290	0.0521	0.0235	0.106	0.0087	0.307	-0.0397	-0.0166
	(0.99)	(1.75)	(-1.19)	(0.31)	(0.09)	(0.65)	(0.13)	(1.40)	(-0.15)	(-0.11)
Non-interest income	0.0026	-0.0292	0.0011	0.0001	-0.0430	-0.0024	-0.0066\	-0.0291	-0.0351	0.0042
	(-0.14)	(-1.22)	(0.68)	(0.01)	(-1.44)	(-0.12)	(-1.47)	(-1.22)	(-1.30)	(0.32)
Loans-to-deposits	-0.119***	-0.0495	-0.0035	-0.123***	-0.0897	-0.113**	-0.0168	-0.0396	-0.0876	-0.115***
	(-2.78)	(-0.74)	(-0.53)	(-3.20)	(-1.15)	(-2.52)	(-0.90)	(-0.62)	(-1.16)	(-2.91)
Tier1 ratio	-0.787**	0.817*	0.0415	-0.784**	0.329	-0.663*	-0.0928	0.932**	0.565	-0.701**
	(-2.40)	(1.80)	(0.79)	(-2.29)	(0.52)	(-1.91)	(-0.88)	(2.11)	(0.92)	(-2.05)
Z-score	0.000458	0.0013	-0.0002	0.0001	0.0012	0.0003	-0.0005	0.0013	0.0010	-0.0001
	(0.43)	(0.69)	(-0.71)	(0.07)	(0.54)	(0.27)	(-1.43)	(0.71)	(0.44)	(-0.07)
Interconnectedness	-0.0165	0.0569	0.0231	-0.0252	-0.0919	0.0002	-0.0011	0.0685	-0.0794	-0.0176
	(-0.17)	(0.42)	(1.30)	(-0.25)	(-0.64)	(0.00)	(-0.04)	(0.49)	(-0.54)	(-0.18)

Table A.5 (cont.)

Unemployment	2.944***	0.197	0.114	2.816***	0.632	2.963***	0.378	0.425	0.375	2.814***
	(3.98)	(0.19)	(0.89)	(4.01)	(0.49)	(3.94)	(1.54)	(0.42)	(0.29)	(4.00)
Inflation	0.840	-1.672	0.161	1.143	-2.476	0.538	0.0829	-1.826	-2.440	0.767
	(0.55)	(-0.74)	(0.62)	(0.69)	(-0.87)	(0.35)	(0.20)	(-0.86)	(-0.87)	(0.49)
GDP growth	1.943**	-1.134	0.224**	2.011**	-1.271	1.870**	0.562*	-1.148	-1.279	2.069**
	(2.19)	(-0.83)	(2.17)	(2.42)	(-0.82)	(2.10)	(1.96)	(-0.85)	(-0.84)	(2.46)
HHI	-6.637***	-0.466	-0.344*	-6.276***	-1.701	-6.603***	-0.755**	-0.693	-1.803	-6.309***
	(-4.77)	(-0.31)	(-1.78)	(-5.02)	(-1.12)	(-4.77)	(-2.13)	(-0.47)	(-1.22)	(-4.98)
Stock market return	0.120	-0.0087	0.0010	0.141*	-0.0894	0.122	0.0174	-0.0237	-0.0689	0.141*
	(1.57)	(-0.09)	(0.10)	(1.84)	(-0.79)	(1.56)	(0.54)	(-0.25)	(-0.65)	(1.83)
Stock market volatility	0.0960	-0.661*	0.0451	-0.0025	-0.0855	0.236	-0.0326	-0.568*	0.0912	0.121
	(0.36)	(-1.90)	(1.14)	(-0.01)	(-0.26)	(0.86)	(-0.50)	(-1.69)	(0.29)	(0.44)
Long-term interest	-0.00443	-0.0015	0.0007	-0.0017	-0.0103	-0.0038	0.0040	-0.0031	-0.0095	-0.0008
	(-0.53)	(-0.11)	(0.63)	(-0.20)	(-0.86)	(-0.46)	(1.62)	(-0.22)	(-0.79)	(-0.09)
Const.	-0.524	3.993**	1.065***	0.315	1.685	-0.652	0.211	3.598*	2.241	0.208
	(-0.35)	(2.01)	(5.38)	(0.22)	(0.75)	(-0.44)	(0.47)	(1.85)	(0.99)	(0.14)
Bank fixed effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Time fixed effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
N	887	860	888	887	859	887	887	860	859	887
R^2	0.134	0.088	0.085	0.138	0.082	0.130	0.108	0.093	0.089	0.138
Adj. R^2	0.102	0.054	0.052	0.106	0.048	0.098	0.076	0.059	0.055	0.107

Dependent variable is the respective rank correlation mentioned at the top of each column. This table contains regression results for rank correlations between exposure-SRMs. A panel approach with fixed effects is applied; t-statistics are given in parentheses. Regressions are performed for each quarter when the value of the VIX volatility index is larger than (or equal to) its 50% quantile of the sample period. The explanatory variables enter the equation with one lag. Regression coefficients are marked with stars if significant as follows: ***= 1% confidence level, **= 5% confidence level, *= 10% confidence level.

Table A.6: Regression results for less volatile market phases

	MES-MES-EQ	MES-SRISK	MES-ExpΔCo-VaR	MES-ExpΔCo-VaR-EQ	MES-EQ-SRISK	MES-EQ - ExpΔCo-VaR	MES-EQ-ExpΔCo-VaR-EQ	SRISK-ExpΔCo-VaR	SRISK-ExpΔCo-VaR EQ	ExpΔCo-VaR-ExpΔCo-VaR-EQ
Assets	0.151**	0.129	0.0239	0.104	0.294**	0.142*	0.0350	0.135	0.294**	0.100
	(2.19)	(1.17)	(1.53)	(1.43)	(2.12)	(1.95)	(1.55)	(1.27)	(2.19)	(1.40)
Reserves-to-loans	1.882**	-1.068	0.0134	1.640**	1.431	1.750**	0.545***	-1.300*	1.370	1.500**
	(2.47)	(-1.54)	(0.15)	(2.07)	(1.59)	(2.38)	(3.20)	(-1.74)	(1.48)	(2.02)
Market-to-book	0.0279	0.147**	-0.0002	0.0121	0.184**	0.0175	0.0228*	0.139**	0.171**	-0.0024
	(0.64)	(2.50)	(-0.04)	(0.28)	(2.44)	(0.40)	(1.75)	(2.30)	(2.36)	(-0.05)
Leverage	1.504**	0.303	0.287***	1.789***	-0.237	1.503**	-0.0892	0.368	0.327	1.676***
	(2.47)	(0.38)	(2.80)	(2.83)	(-0.22)	(2.53)	(-0.39)	(0.45)	(0.32)	(2.65)
ROA	-2.170	-6.216	0.213	-2.791	-6.577	-0.668	0.158	-5.868	-6.749	-1.704
	(-0.55)	(-1.05)	(0.38)	(-0.85)	(-1.24)	(-0.17)	(0.12)	(-0.93)	(-1.27)	(-0.54)
Long-term funding	0.0721	-0.245	-0.0020	0.170	-0.170	0.0998	0.152**	-0.201	-0.108	0.199
	(0.41)	(-0.92)	(-0.03)	(1.01)	(-0.50)	(0.52)	(2.25)	(-0.74)	(-0.31)	(1.13)
Non-interest income	0.0034	-0.0282	0.0022	0.0110	-0.0287*	0.0046	0.0008	-0.0264	-0.0245	0.0145*
	(0.39)	(-1.60)	(0.80)	(1.26)	(-1.85)	(0.52)	(0.22)	(-1.45)	(-1.62)	(1.69)
Loans-to-deposits	-0.0768	0.0009	-0.0059	-0.135**	-0.0450	-0.0771	-0.0463***	-0.0303	-0.0980	-0.137***
	(-1.62)	(0.01)	(-0.61)	(-2.63)	(-0.34)	(-1.61)	(-2.67)	(-0.38)	(-0.76)	(-2.66)
Tier1 ratio	0.784**	0.389	0.213***	0.994***	-0.129	0.750**	0.133	0.526	0.0823	0.947**
	(2.30)	(0.72)	(3.20)	(2.65)	(-0.21)	(2.18)	(1.20)	(0.93)	(0.14)	(2.56)
Z-score	0.0033**	0.0013	-0.0000	0.0027*	0.0058**	0.0030*	0.0004	0.0013	0.00549**	0.0023
	(2.16)	(0.86)	(-0.19)	(1.92)	(2.43)	(1.84)	(0.92)	(0.83)	(2.57)	(1.60)
Interconnectedness	0.0023	-0.113	0.0106	-0.00373	-0.133	-0.0080	-0.0058	-0.113	-0.139	0.0008
	(0.03)	(-0.83)	(0.62)	(-0.05)	(-0.95)	(-0.10)	(-0.23)	(-0.81)	(-1.00)	(0.01)
Unemployment	0.392	2.807***	0.302*	0.708	0.343	0.251	-0.123	2.860***	0.606	0.590
	(0.62)	(2.78)	(1.88)	(1.04)	(0.28)	(0.40)	(-0.71)	(2.88)	(0.51)	(0.87)

Table A.6 (cont.)

Inflation	1.579	3.059	0.366	1.876	-1.072	1.613	0.435	3.196*	-0.970	2.093
	(1.16)	(1.62)	(1.05)	(1.54)	(-0.50)	(1.15)	(0.97)	(1.75)	(-0.42)	(1.65)
GDP growth	-0.995	-1.679	-0.239	-0.960	-2.822	-1.031	0.437	-0.889	-2.783	-0.990
	(-0.97)	(-0.98)	(-1.03)	(-0.95)	(-1.66)	(-1.03)	(1.57)	(-0.52)	(-1.53)	(-1.02)
HHI	-2.440**	-0.268	-0.296	-2.348**	-0.0195	-2.329**	-0.235	-0.281	-0.104	-2.254**
	(-2.48)	(-0.14)	(-1.58)	(-2.24)	(-0.01)	(-2.42)	(-0.82)	(-0.15)	(-0.06)	(-2.19)
Stock market return	0.166*	-0.172	-0.0013	0.143*	-0.0810	0.190**	0.0123	-0.150	-0.0496	0.169**
	(1.91)	(-1.08)	(-0.07)	(1.76)	(-0.40)	(2.11)	(0.44)	(-0.92)	(-0.26)	(2.10)
Stock market volatility	-0.0346	0.0548	0.0664	-0.130	0.172	-0.0464	0.117	-0.0756	0.222	-0.114
	(-0.13)	(0.12)	(1.00)	(-0.46)	(0.44)	(-0.17)	(1.37)	(-0.17)	(0.57)	(-0.41)
Long-term interest	-0.0076	-0.0413***	-0.0006	-0.0034	-0.0115	-0.00804	0.0023	-0.0369***	-0.0148	-0.0054
	(-0.71)	(-3.02)	(-0.41)	(-0.35)	(-0.67)	(-0.73)	(0.96)	(-2.73)	(-0.82)	(-0.54)
Const.	-3.751**	-2.552	0.177	-3.043*	-5.622*	-3.537**	0.253	-2.682	-6.099**	-2.831*
	(-2.50)	(-1.03)	(0.49)	(-1.91)	(-1.76)	(-2.26)	(0.51)	(-1.11)	(-2.01)	(-1.79)
Bank fixed effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Time fixed effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
N	1026	994	1026	1026	994	1026	1026	994	994	1026
R^2	0.080	0.092	0.083	0.087	0.054	0.073	0.087	0.084	0.053	0.083
Adj. R^2	0.049	0.061	0.052	0.056	0.022	0.043	0.057	0.053	0.021	0.052

Dependent variable is the respective rank correlation mentioned at the top of each column. This table contains regression results for rank correlations between exposure-SRMs. A panel approach with fixed effects is applied; t-statistics are given in parentheses. Regressions are performed for each quarter when the value of the VIX volatility index is smaller than (or equal to) its 50% quantile of the sample period. The explanatory variables enter the equation with one lag. Regression coefficients are marked with stars if significant as follows: ***= 1% confidence level, **= 5% confidence level, *= 10% confidence level.

Table A.7: Regression results based on weekly SRMs

	MES-MES-EQ	MES-SRISK	MES-ExpΔCo-VaR	MES-ExpΔCo-VaR-EQ	MES-EQ-SRISK	MES-EQ - ExpΔCo-VaR	MES-EQ-ExpΔCo-VaR-EQ	SRISK-ExpΔCo-VaR	SRISK-ExpΔCo-VaR EQ	ExpΔCo-VaR-ExpΔCo-VaR-EQ
Assets	0.0268	-0.0469	0.0046	0.0264	0.122	-0.0003	0.0252	-0.0290	0.140	-0.0079
	(0.42)	(-0.51)	(0.38)	(0.36)	(1.14)	(-0.00)	(1.40)	(-0.32)	(1.33)	(-0.11)
Reserves-to-loans	0.176	-2.917***	0.0358	0.298	-0.251	-0.0659	0.305	-2.985***	-0.188	0.0367
	(0.18)	(-3.44)	(0.43)	(0.30)	(-0.23)	(-0.07)	(1.65)	(-3.51)	(-0.18)	(0.04)
Market-to-book	0.0105	0.152**	0.0002	0.0107	0.134**	-0.0093	0.0021	0.150**	0.130**	-0.0107
	(0.31)	(2.53)	(0.03)	(0.29)	(2.07)	(-0.26)	(0.19)	(2.64)	(2.03)	(-0.27)
Leverage	-0.278	0.608	-0.0584	-0.332	0.616	-0.308	0.0684	1.022	0.827	-0.518
	(-0.28)	(0.53)	(-0.47)	(-0.38)	(0.43)	(-0.32)	(0.33)	(0.80)	(0.52)	(-0.60)
ROA	4.020	0.898	-0.268	3.997	-5.570	4.289	-0.483	1.145	-5.595	4.447
	(1.12)	(0.11)	(-0.40)	(1.02)	(-0.77)	(1.14)	(-0.49)	(0.13)	(-0.74)	(1.07)
Long-term funding	0.0083	-0.220	-0.0007	-0.0164	-0.156	0.101	0.101*	-0.142	-0.245	0.0447
	(0.05)	(-0.73)	(-0.02)	(-0.10)	(-0.40)	(0.55)	(1.71)	(-0.47)	(-0.61)	(0.25)
Non-interest income	-0.0254	0.0126	-0.0016	-0.0248	0.0061	-0.0256	-0.0055	0.0116	0.0121	-0.0245
	(-1.15)	(0.56)	(-0.81)	(-1.52)	(0.34)	(-1.08)	(-1.29)	(0.56)	(0.77)	(-1.25)
Loans-to-deposits	-0.0548	-0.0118	-0.0062	-0.0847*	-0.0567	-0.0571	-0.0298	-0.0185	-0.0651	-0.0599
	(-1.17)	(-0.14)	(-1.28)	(-1.75)	(-0.46)	(-1.20)	(-1.42)	(-0.23)	(-0.52)	(-1.14)
Tier1 ratio	-0.0016	-0.0002	-0.0001	0.0005	-0.0025	0.0004	-0.0011	0.0004	-0.0008	0.0008
	(-0.31)	(-0.03)	(-0.09)	(0.10)	(-0.37)	(0.09)	(-0.45)	(0.06)	(-0.12)	(0.17)
Z-score	0.0001	0.0007	-0.0002	0.0005	0.0035	-0.0002	0.0003	0.0006	0.0032	0.0002
	(0.06)	(0.36)	(-1.02)	(0.41)	(1.45)	(-0.15)	(0.92)	(0.33)	(1.38)	(0.15)
Interconnectedness	0.0891	0.0201	0.0141	0.0662	-0.178	0.143	0.0645	0.0112	-0.163	0.117
	(0.91)	(0.12)	(0.70)	(0.64)	(-1.07)	(1.35)	(1.63)	(0.07)	(-0.97)	(1.05)
Unemployment	0.691	2.554**	0.199*	0.821	-0.628	0.551	0.173	2.572**	-0.394	0.683
	(0.90)	(2.32)	(1.85)	(1.04)	(-0.52)	(0.69)	(0.79)	(2.29)	(-0.33)	(0.84)

Table A.7 (cont.)

Inflation	-1.148 (-0.86)	-2.410 (-1.07)	0.127 (0.51)	-0.407 (-0.30)	-3.659 (-1.49)	-2.053 (-1.33)	0.315 (1.11)	-2.406 (-1.07)	-2.903 (-1.20)	-1.293 (-0.84)
GDP growth	0.116 (0.11)	1.308 (0.76)	-0.199 (-1.31)	0.204 (0.25)	-0.884 (-0.52)	0.156 (0.15)	0.418* (1.75)	1.530 (0.94)	-0.929 (-0.60)	0.314 (0.35)
HHI	-2.601** (-2.44)	-3.046** (-2.19)	-0.314** (-2.16)	-2.809** (-2.48)	-2.094 (-1.38)	-2.535** (-2.29)	-0.374 (-1.34)	-2.904** (-2.11)	-2.149 (-1.43)	-2.742** (-2.33)
Stock market return	0.224** (2.04)	-0.287* (-1.69)	-0.0197 (-1.30)	0.280*** (2.77)	-0.102 (-0.53)	0.201* (1.86)	-0.0452 (-1.21)	-0.281 (-1.65)	0.0096 (0.05)	0.256*** (2.65)
Stock market volatility	-0.739*** (-2.67)	-0.749* (-1.83)	0.0135 (0.31)	-0.741*** (-3.06)	-0.171 (-0.42)	-0.744** (-2.51)	-0.0530 (-0.77)	-0.790** (-2.01)	-0.191 (-0.49)	-0.762*** (-2.80)
Long-term interest	-0.0078 (-1.04)	-0.0053 (-0.33)	-0.0000 (-0.00)	-0.0084 (-1.07)	-0.0154 (-1.00)	-0.0046 (-0.58)	0.0042* (1.78)	-0.0054 (-0.34)	-0.0173 (-1.11)	-0.0061 (-0.72)
Const.	0.714 (0.47)	1.076 (0.53)	0.940*** (3.48)	0.805 (0.49)	-2.351 (-0.97)	1.224 (0.82)	0.372 (0.77)	0.352 (0.17)	-2.907 (-1.16)	1.610 (1.05)
Bank fixed effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Time fixed effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
N	991	974	990	991	974	990	991	973	974	990
R^2	0.089	0.107	0.090	0.092	0.072	0.088	0.074	0.112	0.067	0.092
Adj. R^2	0.059	0.077	0.060	0.061	0.041	0.058	0.043	0.082	0.036	0.061

Dependent variable is the respective rank correlation mentioned at the top of each column. This table contains regression results for rank correlations between exposure-SRMs based on weekly SRM data. A panel approach with fixed effects is applied; t-statistics are given in parentheses. Regressions are performed on a bi-annual basis, that is bi-annual correlations are computed based on the set of weekly SRMs within half a year. The explanatory variables enter the equation with one lag. Regression coefficients are marked with stars if significant as follows: ***= 1% confidence level, **= 5% confidence level, *= 10% confidence level.

Table A.8: Summary statistics Spearman's cross-sectional rank correlation coefficients

Variable	Mean	Median	Std. Dev.	Min	25%-Quantile	75%-Quantile	Max	Significance
Exposure SRM:								
Corr MES-SRISK	0.2113	0.2 910	0.2889	-0.6020	0.0634	0.4214	0.7094	86% (91%)
Corr MES-MES-EQ	0.1194	0.1914	0.0738	-0.7397	-0.0426	0.3699	0.7823	17% (31%)
Corr MES-ExpΔCo-VaR	0.9571	0.9651	0.0309	0.8103	0.9461	0.9789	0.9921	100%
Corr MES-ExpΔCo-VaR-EQ	0.1226	0.1895	0.3283	-0.7202	-0.0310	0.3654	0.7479	20% (35%)
Corr SRISK-MES-EQ	0.5995	0.6400	0.2341	-0.0869	0.4506	0.7872	0.9962	99% (100%)
Corr SRISK-ExpΔCoVaR	0.27126	0.3238	0.2643	-0.4983	0.1516	0.4539	0.8201	89% (94%)
Corr SRISK-ExpΔCoVaR-EQ	0.6069	0.6502	0.2319	-0.0969	0.4607	0.7902	0.9995	99% (100%)
Corr MES-EQ-ExpΔCoVaR	0.1595	0.1954	0.3227	-0.6953	0.0278	0.3774	0.9165	16% (31%)
Corr MES-EQ-ExpΔCoVaR-EQ	0.9945	0.9964	0.0060	0.9634	0.9940	0.9977	1.0000	100%
Corr ExpΔCoVaR – ExpΔCoVaR-EQ	0.1687	0.2093	0.3197	-0.6688	0.0517	0.3856	0.9112	20% (36%)
Kendall's concordance coefficient								
K_Conc	0.5377	0.5559	0.1449	0.1651	0.4663	0.6472	0.8674	100%

The table shows summary statistics for all correlation pairs. Correlations are computed in a cross-sectional dimension; that is for each day the rank correlations across the whole sample of banks based on a specific pair of SRMs is computed. The column "Significance" indicates the share of correlations that are significantly different from zero at the 99% (95%) level. In order to be used in a cross-sectional regression based on quarterly data, the daily rank correlations are averaged over all days of one quarter. This results in 30 rank correlations for each pair of SRMs. The correlations are computed based on Spearman's correlation coefficient. In addition, summary statistics for Kendall's coefficient of concordance are displayed.

Table A.9: Regression results for cross-sectional rank correlations – USA and Canada

	MES-MES-EQ	**MES-SRISK**	**MES-ExpΔCo-VaR**	**MES-ExpΔCo-VaR-EQ**	**MES-EQ-SRISK**	**MES-EQ-ExpΔCo-VaR**	**MES-EQ-ExpΔCo-VaR-EQ**	**SRISK-ExpΔCo-VaR**	**SRISK-ExpΔCo-VaR EQ**	**ExpΔCo-VaR-ExpΔCo-VaR-EQ**	**K_Conc**
Unemploy-ment	0.824 (1.03)	2.526*** (3.86)	0.109 (1.13)	0.958 (1.17)	2.179* (1.85)	0.956 (1.10)	-0.0108 (-0.91)	2.588*** (3.33)	2.247* (1.86)	1.085 (1.22)	1.224*** (4.68)
Inflation	1.270 (0.74)	3.393* (1.85)	0.334** (2.32)	1.397 (0.80)	1.583 (0.72)	2.501 (1.35)	-0.0478 (-1.41)	3.273 (1.62)	1.489 (0.66)	2.628 (1.40)	1.641** (2.60)
GDP growth	0.847 (0.61)	-3.278* (-1.92)	-0.180 (-1.35)	0.665 (0.47)	-3.329** (-2.16)	0.282 (0.19)	0.0496 (1.54)	-3.934* (-2.05)	-3.386** (-2.15)	0.0402 (0.03)	-0.678 (-1.09)
HHI	-5.581 (-0.36)	15.39 (0.82)	-0.719 (-0.57)	-2.031 (-0.13)	5.112 (0.26)	-6.466 (-0.38)	-0.0878 (-0.44)	17.65 (0.87)	6.791 (0.34)	-2.943 (-0.17)	6.047 (0.72)
Stock market return	-0.335* (-2.04)	0.124 (0.77)	-0.0036 (-0.30)	-0.331* (-1.97)	0.265 (1.61)	-0.346* (-1.97)	0.0005 (0.20)	0.134 (0.79)	0.252 (1.51)	-0.340* (-1.90)	-0.0599 (-0.94)
Stock market volatility	-1.498** (-2.82)	-0.803** (-2.22)	0.0465 (1.57)	-1.538** (-2.80)	2.788*** (7.85)	-1.591** (-2.82)	0.0114 (1.64)	-0.542 (-1.46)	2.778*** (7.72)	-1.654*** (-2.85)	-0.208 (-1.02)
Long-term in-terest	-0.0295 (-0.87)	-0.105*** (-3.03)	-0.00415 (-1.69)	-0.0317 (-0.95)	-0.131*** (-4.33)	-0.0346 (-0.93)	0.0002 (0.22)	-0.121*** (-3.06)	-0.133*** (-4.18)	-0.0363 (-0.99)	-0.0488*** (-4.48)
Const.	0.794 (0.81)	-0.506 (-0.40)	1.013*** (12.36)	0.547 (0.55)	-0.372 (-0.27)	0.878 (0.80)	1.001*** (76.55)	-0.672 (-0.49)	-0.479 (-0.35)	0.643 (0.58)	0.140 (0.26)
N	29	29	29	29	29	29	29	29	29	29	29
R^2	0.585	0.479	0.324	0.565	0.850	0.243	0.243	0.456	0.843	0.534	0.559
Adj. R^2	0.447	0.305	0.099	0.420	0.800	-0.010	-0.010	0.274	0.790	0.379	0.412

Dependent variable is the (cross-sectional) rank correlation between the various exposure SRMs based on Spearman's correlation coefficient and Kendall's coefficient of concordance, respectively, with t-statistics given in parentheses. An OLS approach with heteroscedasticity-robust standard errors is applied. Regression coefficients are marked with stars if significant as follows: ***= 1% confidence level, **= 5% confidence level, *= 10% confidence level.

Table A.10: Regression results for cross-sectional rank correlations – UK and Ireland

	MES-MES-EQ	MES-SRISK	MES-ExpΔCo-VaR	MES-ExpΔCo-VaR-EQ	MES-EQ-SRISK	MES-EQ-ExpΔCo-VaR	MES-EQ-ExpΔCo-VaR-EQ	SRISK-ExpΔCo-VaR	SRISK-ExpΔCo-VaR EQ	ExpΔCo-VaR-ExpΔCo-VaR-EQ	K_Conc
Unemployment	-2.363*** (-3.01)	-1.447 (-1.09)	0.318* (1.84)	-2.291*** (-2.85)	1.368 (1.54)	-1.972** (-2.42)	0.166** (2.74)	-2.062 (-1.62)	0.668 (0.82)	-2.234** (-2.70)	-0.831 (-1.45)
Inflation	-0.354 (-0.05)	2.899 (0.64)	-0.984 (-1.64)	-0.667 (-0.09)	4.053 (1.46)	1.774 (0.23)	0.0517 (0.49)	7.458 (1.51)	3.685 (1.43)	1.312 (0.17)	0.825 (0.24)
GDP growth	1.646 (0.78)	-0.171 (-0.11)	0.0016 (0.00)	1.764 (0.85)	0.0823 (0.07)	1.664 (0.72)	-0.130** (-2.17)	-0.375 (-0.22)	0.489 (0.46)	2.012 (0.87)	0.346 (0.32)
HHI	-12.34 (-0.43)	-0.855 (-0.03)	0.527 (0.11)	-12.27 (-0.42)	1.079 (0.06)	-8.118 (-0.27)	0.688 (1.04)	1.228 (0.05)	-0.786 (-0.05)	-8.487 (-0.28)	-3.552 (-0.23)
Stock market return	-0.532 (-1.05)	-0.127 (-0.29)	0.0039 (0.11)	-0.519 (-1.02)	0.0817 (0.39)	-0.464 (-0.89)	0.0103 (1.44)	-0.158 (-0.34)	0.0225 (0.11)	-0.474 (-0.90)	-0.118 (-0.44)
Stock market volatility	0.880 (0.83)	2.191* (1.90)	0.353** (2.70)	1.057 (0.99)	1.641** (2.46)	0.892 (0.82)	0.0404** (2.11)	1.491 (1.19)	1.465** (2.32)	0.963 (0.84)	0.863 (1.56)
Long-term interest	-0.0770 (-1.29)	-0.0686 (-1.43)	0.0032 (0.38)	-0.0805 (-1.35)	-0.123*** (-4.48)	-0.0922 (-1.43)	0.0011 (0.82)	-0.105* (-2.00)	-0.125*** (-5.11)	-0.096 (-1.48)	-0.0546* (-1.81)
Const.	0.748 (0.47)	-0.0108 (-0.01)	0.800*** (2.99)	0.716 (0.45)	0.323 (0.32)	0.589 (0.36)	0.926*** (23.93)	0.283 (0.18)	0.559 (0.61)	0.648 (0.39)	0.676 (0.79)
N	29	29	29	29	29	29	29	29	29	29	29
R^2	0.481	0.316	0.397	0.499	0.524	0.414	0.662	0.240	0.615	0.453	0.403
Adj. R^2	0.309	0.088	0.196	0.332	0.365	0.219	0.550	-0.013	0.487	0.271	0.204

Dependent variable is the (cross-sectional) rank correlation between the various exposure SRMs based on Spearman's correlation coefficient and Kendall's coefficient of concordance, respectively, with t-statistics given in parentheses. An OLS approach with heteroscedasticity-robust standard errors is applied. Regression coefficients are marked with stars if significant as follows: ***= 1% confidence level, **= 5% confidence level, *= 10% confidence level.

Table A.11: Regression results for cross-sectional rank correlations – Germany, Switzerland, and Austria

	MES-MES-EQ	**MES-SRISK**	**MES-ExpΔCo-VaR**	**MES-ExpΔCo-VaR-EQ**	**MES-EQ-SRISK**	**MES-EQ-ExpΔCo-VaR**	**MES-EQ-ExpΔCo-VaR-EQ**	**SRISK-ExpΔCo-VaR**	**SRISK-ExpΔCo-VaR EQ**	**ExpΔCo-VaR-ExpΔCo-VaR-EQ**	**K_Conc**
Unemploy-ment	-1.035 (-0.73)	0.739 (0.58)	-0.0505 (-0.58)	-1.132 (-0.79)	0.473 (1.38)	0.473 (1.38)	0.0131 (0.38)	1.049 (0.84)	0.458 (1.64)	-0.899 (-0.63)	-0.117 (-0.17)
Inflation	-12.38 (-0.94)	-10.17 (-0.92)	0.0697 (0.09)	-12.96 (-0.96)	24.37*** (6.80)	24.37*** (6.80)	0.210 (0.67)	-10.40 (-0.97)	23.64*** (6.56)	-7.971 (-0.59)	-1.427 (-0.23)
GDP growth	1.261 (0.33)	-1.696 (-0.63)	0.271 (1.19)	1.482 (0.39)	-3.772*** (-4.08)	-3.772*** (-4.08)	0.126 (1.13)	-1.461 (-0.55)	-3.627*** (-3.63)	0.215 (0.06)	-0.424 (-0.27)
HHI	-65.56*** (-2.87)	-70.05*** (-3.56)	1.196 (1.09)	-64.84** (-2.82)	-7.923 (-1.41)	-7.923 (-1.41)	0.158 (0.34)	-66.87*** (-3.66)	-8.163 (-1.45)	-62.37*** (-2.88)	-32.21*** (-3.05)
Stock market return	0.495 (1.64)	0.358 (1.32)	0.00722 (0.44)	0.468 (1.54)	0.0141 (0.20)	0.456 (1.53)	0.0029 (0.43)	0.323 (1.26)	0.0121 (0.18)	0.435 (1.45)	0.211 (1.52)
Stock market volatility	1.557** (2.58)	1.272** (2.48)	0.0539 (1.11)	1.515** (2.48)	0.962*** (5.55)	0.962*** (5.55)	-0.0095 (-0.72)	0.885* (1.81)	0.963*** (5.42)	1.305** (2.14)	0.781*** (2.83)
Long-term in-terest	0.359* (1.88)	0.395** (2.40)	-0.0152* (-1.99)	0.360* (1.89)	-0.0762* (-1.83)	-0.0762* (-1.83)	-0.0049 (-1.13)	0.389** (2.54)	-0.0700 (-1.64)	0.339* (1.90)	0.160* (1.85)
Const.	2.709** (2.78)	2.983*** (3.61)	0.894*** (15.98)	2.701** (2.75)	0.917*** (3.54)	0.917***(4 8.94)	0.988*** (48.94)	2.960*** (3.84)	0.938*** (3.64)	2.626** (2.82)	1.812*** (4.01)
N	29	29	29	29	29	29	29	29	29	29	29
R^2	0.461	0.564	0.334	0.447	0.782	0.782	0.439	0.577	0.777	0.424	0.459
Adj. R^2	0.281	0.419	0.112	0.263	0.709	0.709	0.252	0.436	0.703	0.232	0.278

Dependent variable is the (cross-sectional) rank correlation between the various exposure SRMs based on Spearman's correlation coefficient and Kendall's coefficient of concordance, respectively, with t-statistics given in parentheses. An OLS approach with heteroscedasticity-robust standard errors is applied. Regression coefficients are marked with stars if significant as follows: ***= 1% confidence level, **= 5% confidence level, *= 10% confidence level.

Table A.12: Regression results for cross-sectional rank correlations – France, Netherlands and Belgium

	MES-MES-EQ	MES-SRISK	MES-ExpΔCo-VaR	MES-ExpΔCo-VaR-EQ	MES-EQ-SRISK	MES-EQ-ExpΔCo-VaR	MES-EQ-ExpΔCo-VaR-EQ	SRISK-ExpΔCo-VaR	SRISK-ExpΔCo-VaR EQ	ExpΔCo-VaR-ExpΔCo-VaR-EQ	K_Conc
Unemployment	2.329*** (4.26)	1.263*** (3.08)	0.319*** (2.93)	2.328*** (4.24)	-0.280 (-1.42)	1.674*** (3.05)	-0.0220 (-1.70)	0.645 (1.71)	-0.290 (-1.50)	1.676*** (3.01)	0.771*** (3.69)
Inflation	8.895 (1.22)	8.316 (1.11)	0.480 (0.44)	9.161 (1.26)	-2.590 (-1.27)	6.967 (0.99)	-0.0663 (-0.56)	4.489 (0.62)	-3.039 (-1.42)	7.432 (1.04)	3.204 (1.08)
GDP growth	-2.715 (-0.83)	-0.428 (-0.13)	-0.0144 (-0.03)	-2.378 (-0.72)	0.912 (1.06)	-3.044 (-0.98)	0.0538 (1.19)	0.422 (0.13)	1.069 (1.18)	-2.825 (-0.89)	-0.716 (-0.54)
HHI	66.05** (2.22)	53.13 (1.41)	-2.048 (-0.37)	64.53** (2.19)	-9.972 (-1.62)	60.03** (2.22)	-0.280 (-0.63)	40.92 (1.16)	-10.56* (-1.73)	59.23** (2.21)	25.68* (1.88)
Stock market return	-0.345 (-1.04)	-0.0997 (-0.24)	-0.0349 (-0.96)	-0.361 (-1.09)	0.0228 (0.46)	-0.359 (-1.14)	-0.0050 (-1.16)	-0.137 (-0.33)	0.0286 (0.59)	-0.362 (-1.14)	-0.132 (-0.81)
Stock market volatility	-0.697 (-1.23)	0.0459 (0.06)	0.0947 (1.03)	-0.725 (-1.27)	0.380** (2.41)	-0.878 (-1.61)	-0.0022 (-0.21)	-0.178 (-0.23)	0.393** (2.65)	-0.901 (-1.64)	-0.197 (-0.72)
Long-term interest	-0.179*** (-4.76)	-0.226*** (-5.40)	-0.0086 (-1.12)	-0.173*** (-4.57)	0.0114 (1.03)	-0.157*** (-4.24)	0.0011* (1.92)	-0.191*** (-4.26)	0.0136 (1.27)	-0.153*** (-4.05)	-0.0848*** (-5.08)
Const.	-8.229* (-2.05)	-6.574 (-1.29)	1.169 (1.58)	-8.043* (-2.02)	2.075** (2.56)	-7.232* (-1.98)	1.032*** (17.17)	-4.758 (-0.99)	2.148** (2.66)	-7.136* (-1.97)	-2.644 (-1.43)
N	29	29	29	29	29	29	29	29	29	29	29
R^2	0.733	0.701	0.170	0.727	0.464	0.719	0.219	0.661	0.507	0.707	0.731
Adj. R^2	0.645	0.601	-0.107	0.636	0.286	0.625	-0.041	0.548	0.342	0.610	0.641

Dependent variable is the (cross-sectional) rank correlation between the various exposure SRMs based on Spearman's correlation coefficient and Kendall's coefficient of concordance, respectively, with t-statistics given in parentheses. An OLS approach with heteroscedasticity-robust standard errors is applied. Regression coefficients are marked with stars if significant as follows: ***= 1% confidence level, **= 5% confidence level, *= 10% confidence level.

Table A.13: Regression results for cross-sectional rank correlations – Portugal, Spain, Italy and Greece

	MES-MES-EQ	MES-SRISK	MES-ExpΔCo-VaR	MES-ExpΔCo-VaR-EQ	MES-EQ-SRISK	MES-EQ-ExpΔCo-VaR	MES-EQ-ExpΔCo-VaR-EQ	SRISK-ExpΔCo-VaR	SRISK-ExpΔCo-VaR EQ	ExpΔCo-VaR-ExpΔCo-VaR-EQ	K_Conc
Unemployment	0.877 (0.66)	0.0725 (0.10)	-0.0402* (-1.92)	0.912 (0.67)	0.0029 (0.01)	0.651 (0.49)	-0.0210*** (-4.84)	-0.0338 (-0.04)	0.0551 (0.20)	0.711 (0.53)	0.240 (0.46)
Inflation	-7.929 (-0.53)	-1.484 (-0.18)	0.477* (1.74)	-7.875 (-0.52)	5.112 (1.27)	-7.439 (-0.49)	0.0584 (0.71)	-1.995 (-0.24)	4.828 (1.21)	-7.657 (-0.50)	-2.320 (-0.39)
GDP growth	2.748 (0.95)	3.714** (2.58)	-0.0747 (-1.18)	2.854 (0.98)	0.518 (0.69)	2.478 (0.85)	-0.0552*** (-2.83)	3.487** (2.48)	0.549 (0.74)	2.627 (0.90)	1.384 (1.23)
HHI	-9.321 (-1.66)	-8.265** (-2.81)	0.194* (1.75)	-9.487 (-1.67)	-3.030** (-2.13)	-8.478 (-1.50)	0.0406 (1.51)	-8.658*** (-2.98)	-3.250** (-2.30)	-8.720 (-1.53)	-4.951** (-2.23)
Stock market return	0.0450 (0.28)	-0.0652 (-0.55)	-0.0020 (-0.40)	0.0500 (0.31)	-0.0960 (-1.09)	0.0391 (0.24)	-0.0014 (-0.95)	-0.0681 (-0.56)	-0.0979 (-1.11)	0.0431 (0.26)	-0.0170 (-0.28)
Stock market volatility	-0.0178 (-0.03)	0.205 (0.56)	0.0360* (1.87)	-0.0453 (-0.07)	0.306 (1.01)	0.0422 (0.07)	0.0034 (0.64)	0.113 (0.30)	0.272 (0.88)	0.0032 (0.00)	0.0628 (0.27)
Long-term interest	0.0490 (0.55)	0.0002 (0.00)	-0.0019 (-1.08)	0.0484 (0.53)	-0.0147 (-0.58)	0.0507 (0.56)	-0.0002 (-0.43)	0.0011 (0.02)	-0.0133 (-0.52)	0.0516 (0.56)	0.0144 (0.41)
Const.	0.872** (2.45)	1.031*** (4.77)	0.967*** (82.71)	0.895** (2.48)	0.864*** (5.71)	0.766** (2.14)	0.996*** (298.65)	1.094*** (5.03)	0.882*** (5.82)	0.793** (2.18)	0.959*** (6.74)
N	29	29	29	29	29	29	29	29	29	29	29
R^2	0.314	0.540	0.518	0.309	0.766	0.306	0.576	0.546	0.770	0.300	0.518
Adj. R^2	0.085	0.387	0.357	0.079	0.688	0.075	0.435	0.395	0.693	0.066	0.357

Dependent variable is the (cross-sectional) rank correlation between the various exposure SRMs based on Spearman's correlation coefficient and Kendall's coefficient of concordance, respectively, with t-statistics given in parentheses. An OLS approach with heteroscedasticity-robust standard errors is applied. Regression coefficients are marked with stars if significant as follows: ***= 1% confidence level, **= 5% confidence level, *= 10% confidence level.

Table A.14: Regression results for cross-sectional rank correlations – Norway, Sweden and Denmark

	MES-MES-EQ	MES-SRISK	MES-ExpΔCo-VaR	MES-ExpΔCo-VaR-EQ	MES-EQ-SRISK	MES-EQ-ExpΔCo-VaR	MES-EQ-ExpΔCo-VaR-EQ	SRISK-ExpΔCo-VaR	SRISK-ExpΔCo-VaR EQ	ExpΔCo-VaR-ExpΔCo-VaR-EQ	K_Conc
Unemployment	2.381** (2.77)	1.562* (2.03)	-0.0302 (-0.41)	2.231** (2.51)	-0.253* (-1.94)	2.299** (2.51)	-0.0683*** (-5.79)	1.375* (2.02)	-0.272 (-1.52)	2.203** (2.35)	0.914*** (2.94)
Inflation	6.191* (2.07)	1.209 (0.34)	-0.377 (-0.74)	7.069** (2.33)	-1.286 (-0.67)	7.524** (2.41)	0.265** (2.59)	3.130 (0.91)	-1.813 (-0.92)	8.345** (2.70)	2.421* (1.75)
GDP growth	-2.732** (-2.20)	-0.894 (-0.70)	0.222 (1.42)	-2.789** (-2.21)	1.140 (1.31)	-3.330** (-2.65)	-0.0214 (-0.56)	-1.338 (-1.04)	1.295 (1.44)	-3.409** (-2.71)	-0.949 (-1.58)
HHI	-7.044 (-0.62)	-33.05*** (-2.97)	-2.691** (-2.17)	-6.443 (-0.56)	-20.85** (-2.74)	-0.539 (-0.05)	-0.587** (-2.82)	-28.62** (-2.64)	-22.34** (-2.81)	0.473 (0.04)	-9.735* (-1.74)
Stock market return	0.0184 (0.14)	-0.0390 (-0.25)	-0.0059 (-0.35)	0.0075 (0.06)	0.0315 (0.45)	0.0329 (0.25)	-0.0055 (-1.55)	-0.0258 (-0.18)	0.0346 (0.47)	0.0219 (0.17)	0.0056 (0.09)
Stock market volatility	-0.298 (-1.06)	-0.396 (-1.23)	0.100*** (2.91)	-0.352 (-1.21)	0.292* (1.86)	-0.405 (-1.42)	-0.0108 (-1.48)	-0.543 (-1.71)	0.310* (1.90)	-0.467 (-1.63)	-0.142 (-1.09)
Long-term interest	0.0318 (1.30)	0.0244 (0.83)	-0.0016 (-0.46)	0.0374 (1.52)	-0.0362* (-2.06)	0.0287 (1.13)	0.0005 (0.99)	0.0213 (0.74)	-0.0350* (-1.91)	0.0341 (1.35)	0.0084 (0.70)
Const.	1.021 (0.72)	4.235*** (3.07)	1.278*** (8.14)	0.949 (0.66)	3.360*** (3.53)	0.295 (0.22)	1.069*** (42.02)	3.780** (2.82)	3.543*** (3.56)	0.174 (0.13)	1.776** (2.55)
N	29	29	29	29	29	29	29	29	29	29	29
R^2	0.528	0.466	0.405	0.540	0.504	0.498	0.736	0.478	0.500	0.515	0.529
Adj. R^2	0.370	0.289	0.206	0.387	0.338	0.331	0.648	0.304	0.333	0.354	0.372

Dependent variable is the (cross-sectional) rank correlation between the various exposure SRMs based on Spearman's correlation coefficient and Kendall's coefficient of concordance, respectively, with t-statistics given in parentheses. An OLS approach with heteroscedasticity-robust standard errors is applied. Regression coefficients are marked with stars if significant as follows: ***= 1% confidence level, **= 5% confidence level, *= 10% confidence level.

Appendix B: (Essay 2)

Table B.1: Samp1e constituents

Cashtag	Group	Name	Country
0939	G-SIFI	China Construction Bank	CN
1288	G-SIFI	Agricultural Bank of China	CN
1398	G-SIFI	Industrial and Commercial Bank of China	CN
3988	G-SIFI	Bank of China	CN
AAPL	S&P100	Apple	US
ABBV	S&P100	AbbVie	US
ABT	S&P100	Abbott Laboratories	US
ACA	G-SIFI	Credit Agricole	FR
ACN	S&P100	Accenture	US
AIG	S&P100	AIG	US
ALL	S&P100	Allstate	US
ALLY	LD-SIFI	Ally Financial	US
AMGN	S&P100	Amgen	US
AMZN	S&P100	Amazon	US
ANZ	LD-SIFI	Australian New Zealand Banking	AUS
APA	S&P100	Apache	US
APC	S&P100	Anadarko Petroleum	US
AXP	S&P100, LD-SIFI	American Express	US
BA	S&P100	Boeing	US
BAC	S&P100, G-SIFI	Bank of America	US
BAX	S&P100	Baxter International	US
BBT	LD-SIFI	BB&T	US
BBVA	LD-SIFI	BBVA	ES
BCS	G-SIFI	Barclays	UK
BIIB	S&P100	Biogen	US
BK	S&P100, G-SIFI	Bank of New York Mellon	US
BLK	S&P100	BlackRock	US
BMO	LD-SIFI	Bank of Montreal	CA
BMY	S&P100	Bristol-Myers Squibb	US
BNP	G-SIFI	BNP Paribas	FR
BNS	LD-SIFI	Scotiabank	CA
BRK.B	S&P100	Berkshire Hathaway	US
C	S&P100	Citigroup	US
CAC40	Index	CAC40	FR
CAT	S&P100	Caterpillar	US
CBA	LD-SIFI	Commonwealth Bank of AUS	AUS
CL	S&P100	Colgate-Palmolive	US
CM	LD-SIFI	Canadian Imperial Bank of Commerce	CA
CMA	LD-SIFI	Comerica	US

Table B.1 (cont.)

CMCSA	S&P100	Comcast	US
COF	S&P100, LD-SIFI	Capital One Financial	US
COP	S&P100	ConocoPhillips	US
COST	S&P100	Costco	US
CS	G-SIFI	Credit Suisse	CH
CSCO	S&P100	Cisco Systems	US
CVS	S&P100	CVS Caremark	US
CVX	S&P100	Chevron	US
DANSKE	LD-SIFI	Dankse Bank	DK
DAX	Index	DAX	DE
DB	G-SIFI	Deutsche Bank	DE
DD	S&P100	DuPont	US
DFS	LD-SIFI	Discover Financial Services	US
DIS	S&P100	Walt Disney	US
DNBO	LD-SIFI	DNB Bank	NO
DOW	S&P100	Dow Chemical	US
DVN	S&P100	Devon Energy	US
EBAY	S&P100	eBay	US
EMC	S&P100	EMC	US
EMR	S&P100	Emerson	US
EXC	S&P100	Exelon	US
F	S&P100	Ford	US
FB	S&P100	Facebook	US
FCX	S&P100	Freeport-McMoran	US
FDX	S&P100	FedEx	US
FITB	LD-SIFI	FifthThird Bank	US
FOXA	S&P100	21Century Fox, Inc.	US
FTSE	Index	FTSE100	UK
GD	S&P100	General Dynamics	US
GE	S&P100	General Electric	US
GILD	S&P100	Gilead	US
GLE	G-SIFI	Societe Generale	FR
GM	S&P100	General Motors	US
GOOG	S&P100	Google	US
GS	S&P100, G-SIFI	Goldman Sachs	US
HAL	S&P100	Halliburton	US
HBAN	LD-SIFI	Huntington Bancshares	US
HD	S&P100	Home Depot	US
HON	S&P100	Honeywell	US
HSBC	G-SIFI	HSBC	UK
IBM	S&P100	IBM	US
ING	G-SIFI	ING Group	NL
INTC	S&P100	Intel	US

Table B.1 (cont.)

JNJ	S&P100	Johnson & Johnson	US
JPM	S&P100, G-SIFI	JP Morgan Chase	US
JYSK	LD-SIFI	Jyske Bank	DK
KEY	LD-SIFI	Keybank	US
KO	S&P100	Coca-Cola	US
LLY	S&P100	Eli Lilly	US
LMT	S&P100	Lockheed-Martin	US
LOW	S&P100	Lowe's	US
LYG	LD-SIFI	Llyd's Banking Group	UK
MA	S&P100	Mastercard Inc.	US
MCD	S&P100	McDonalds Corp	US
MDLZ	S&P100	Mondelez	US
MDT	S&P100	Medtronic	US
MET	S&P100, LD-SIFI	Metlife	US
MFG	G-SIFI	Mizuho	JP
MO	S&P100	Altria	US
MON	S&P100	Monsanto	US
MRK	S&P100	Merck	US
MS	S&P100, G-SIFI	Morgan Stanley	US
MSFT	S&P100	Microsoft	US
MTB	LD-SIFI	M&T Bank	US
MTU	G-SIFI	Bank of Tokyo-Mitsubishi	JP
NAB	LD-SIFI	National Australian Bank	AUS
NDX1	G-SIFI	Nordea	FI
NKE	S&P100	Nike	US
NOV	S&P100	National Oilwell	US
NSC	S&P100	Norfolk Southern	US
NTRS	LD-SIFI	NorthernTrust	US
ORCL	S&P100	Oracle	US
OXY	S&P100	Occidental Petroleum	US
PEP	S&P100	Pepsico	US
PFE	S&P100	Pfizer	US
PG	S&P100	Procter & Gamble	US
PM	S&P100	Phillip Morris	US
PNC	LD-SIFI	PNC Financial Services	US
QCOM	S&P100	Qualcomm	US
RBS	G-SIFI	Royal Bank of Scotland	UK
RF	LD-SIFI	RegionsFinancial	US
RTN	S&P100	Raytheon	US
RY	LD-SIFI	Royal Bank of Canada	CA
SAN	G-SIFI	Satander Group	ES
SBUX	S&P100	Starbucks	US
SEBA	LD-SIFI	SEB	SE

Table B.1 (cont.)

SLB	S&P100	Schlumberger	US
SMFG	G-SIFI	Sumitomo Mitsui	JP
SO	S&P100	Southern Company	US
SPG	S&P100	Simon Property	US
SPX	Index	S&P500	US
STAN	G-SIFI	Standard Chartered	UK
STI	LD-SIFI	SunTrust Banks	US
STOXX	Index	EuroStoxx50	EU
STT	G-SIFI	State Street	US
SYDB	LD-SIFI	Sydbank	DK
T	S&P100	AT&T	US
TD	LD-SIFI	Toronto Dominion Bank	CA
TGT	S&P100	Target	US
TWX	S&P100	Time Warner	US
TXN	S&P100	Texas Instruments	US
UBS	G-SIFI	UBS	CH
UCGR	G-SIFI	UniCredit	IT
UNH	S&P100	United Health	US
UNP	S&P100	Union Pacific	US
UPS	S&P100	UPS	US
USB	S&P100, LD-SIFI	US Bancorp	US
UTX	S&P100	United Technologies	US
V	S&P100	Visa	US
VZ	S&P100	Verizon	US
WBA	S&P100	Walgreens Boots Alliance	US
WFC	S&P100, G-SIFI	Wells Fargo	US
WMT	S&P100	Walmart	US
WPC	LD-SIFI	Westpac Banking	AUS
XOM	S&P100	Exxon Mobil	US
ZION	LD-SIFI	Zions Bancorp	US

Cashtag represents the ticker symbol that is used in conjunction with the $-sign and name refers to the keyword, respectively, that is used to collect tweets. Group indicates to which group the stock belongs, that is either G-SIFI for globally systemically important financial institution, LD-SIFI for local/domestic systemically important financial institution and/or constituent of the S&P100 index. Naturally, the latter are exclusively from the US while the others are globally distributed.

Table B.2: Out-of-sample forecast evaluation of HAR-model -p-values of Clark-West-Test

		No jumps				Jumps			
		Mean	**Min**	**Median**	**Max**	**Mean**	**Min.**	**Median**	**Max.**
Overall									
	Cashtag- BM								
	RV^2	0.4848	0.0000	0.4621	0.9999	0.4581	0.0000	0.3484	1.0000
	RV	0.4824	0.0000	0.4796	0.9978	0.4616	0.0000	0.3958	0.9995
	logRV	0.4802	0.0012	0.4698	0.9982				
	Company name - BM								
	RV^2	0.4822	0.0000	0.4737	1.0000	0.478	0.0000	0.4832	1.0000
	RV	0.4723	0.0000	0.4659	0.9999	0.4834	0.0000	0.4730	1.0000
	logRV	0.4360	0.0001	0.3720	0.9994				
SP100									
	Cashtag - BM								
	RV^2	0.4888	0.0000	0.4888	0.9890	0.4263	0.0000	0.2972	0.9864
	RV	0.5049	0.0000	0.5214	0.9664	0.4508	0.0000	0.3919	0.9822
	logRV	0.5051	0.0248	0.5061	0.9982				
	Company name - BM								
	RV^2	0.4955	0.0000	0.4672	1.0000	0.4895	0.0000	0.4797	1.0000
	RV	0.4945	0.0000	0.4825	0.9993	0.4704	0.0000	0.4405	0.9994
	logRV	0.4375	0.0001	0.3696	0.9793				
SIFI									
	Cashtag - BM								
	RV^2	0.4670	0.0000	0.3933	0.9999	0.4909	0.0000	0.5051	1.0000
	RV	0.4389	0.0000	0.3505	0.9978	0.4827	0.0000	0.5112	0.9995
	logRV	0.4365	0.0012	0.3062	0.9962				
	Company name - BM								
	RV^2	0.4478	0.0000	0.4750	1.0000	0.4451	0.0000	0.3608	1.0000
	RV	0.4945	0.0000	0.4825	0.9999	0.5046	0.0032	0.5133	1.0000
	logRV	0.4478	0.0010	0.3994	0.9994				
Index									
	Cashtag - BM								
	RV^2	0.6125	0.2506	0.6715	0.8571	0.7724	0.6470	0.7667	0.9094
	RV	0.5076	0.1775	0.5099	0.8331	0.4467	0.0564	0.4460	0.8383
	logRV	0.4559	0.0893	0.4469	0.8405				
	Company name - BM								
	RV^2	0.6050	0.2559	0.6127	0.9385	0.6414	0.1912	0.7422	0.8899
	RV	0.5410	0.1918	0.4964	0.9794	0.5189	0.0695	0.5122	0.9818
	logRV	0.2590	0.0078	0.1143	0.7994				

This table displays summary statistics of p-values resulting from the Clark-West-Test to test for significant differences between MSEs of the benchmark and the respective Twitter model. Significance can be for instance rejected at the 5% level for p>0.05.

Technical Appendix A (Essay1)

We strictly follow Benoit et al. (2013) and Benoit et al. (2017) with respect to the computation of the three SRMs MES, SRISK and ΔCoVaR. We refer to their work and the references therein for more details of the statistical and econometrical issues involved.

Benoit et al. (2013) note that there is no unique way to compute the respective SRMs. More precisely, the original versions are outlined with very different econometric methods which could lead to problems when they are jointly analyzed since results might be distorted by model risk. They propose a unified multivariate DCC-GARCH framework where all three major SRMs are computed based on some few common parameters.

This framework builds upon work from Brownlees and Engle (2012) where the two-dimensional vector r_t made of the demeaned market equity return $r_{m,t}$ and demeaned bank equity return $r_{i,t}$ is defined as (see Brownlees and Engle (2012), p. 9 and Benoit et al. (2017), p. 136)

$$r_t = \sqrt{H_t}\nu_t \tag{A1}$$

with the conditional variance-covariance matrix H_t being defined as (Benoit et al. (2017), p. 137)

$$H_t = \begin{pmatrix} \sigma_{m,t}^2 & \sigma_{i,t}\sigma_{m,t}\rho_{i,t} \\ \sigma_{i,t}\sigma_{m,t}\rho_{i,t} & \sigma_{i,t}^2 \end{pmatrix}. \tag{A2}$$

The vector $\nu_t' = (\varepsilon_{m,t} \quad \xi_{i,t})$ is supposed to be i.i.d. and standard normally distributed with $E(\nu_t) = 0$ and $E(\nu_t\nu_t') = I_2$. Estimates for the conditional market equity return variance $\sigma_{m,t}^2$, the conditional bank equity return variance $\sigma_{i,t}^2$ as well as the conditional correlation between market and bank equity returns $\rho_{i,t}$ are in the following employed as main ingredients for the computation of the SRMs.

Marginal expected shortfall (MES)

In equation (2.1) in the main text, MES is defined as

$$MES_{it}(\tau) = -E_{t-1}\left(r_{i,t} \middle| r_{m,t} < \tau\right). \tag{A3}$$

MES is multiplied with (-1) such that larger values represent larger systemic risk. Benoit et al. (2013, p.10) show that in the presence of the realistic assumption of nonlinear dependencies between market and bank equity returns this representation can be estimated as

$$\widehat{MES}_{i,t}(\tau) = -\hat{\sigma}_{i,t}\hat{\rho}_{i,t}E_{t-1}\left(\varepsilon_{m,t}\middle|\varepsilon_{m,t} < \kappa\right) - \hat{\sigma}_{i,t}\sqrt{1-\hat{\rho}_{i,t}{}^2}E_{t-1}\left(\xi_{i,t}\middle|\varepsilon_{m,t} < \kappa\right). \tag{A4}$$

The conditional variance $\sigma_{i,t}^2$ is estimated in the framework of a GJR-GARCH (1,1) specification with a pseudo maximum likelihood approach similar to the conditional correlation $\rho_{i,t}$ which is estimated in a dynamic conditional correlations model based on the work of Engle (2002). Equation (A4) explicitly takes nonlinear dependencies between market and bank returns into account that are not captured by the (linear) correlation $\rho_{i,t}$.This requires to estimate the conditional tail expectation according to (see Brownlees and Engle (2012), p.15)

$$E_{t-1}\left(\xi_{i,t}\middle|\varepsilon_{m,t} < \kappa\right) = \frac{\sum_{t=1}^{T} K\left(\frac{\kappa - \varepsilon_{m,t}}{\lambda}\right) \cdot \xi_{i,t}}{\sum_{t=1}^{T} K\left(\frac{\kappa - \varepsilon_{m,t}}{\lambda}\right)} \tag{A5}$$

and

$$E_{t-1}\left(\varepsilon_{m,t}\middle|\varepsilon_{m,t} < \kappa\right) = \frac{\sum_{t=1}^{T} K\left(\frac{\kappa - \varepsilon_{m,t}}{\lambda}\right) \cdot \varepsilon_{m,t}}{\sum_{t=1}^{T} K\left(\frac{\kappa - \varepsilon_{m,t}}{\lambda}\right)}. \tag{A6}$$

Brownlees and Engel (2012) suggest to use a nonparametric kernel density approach to estimate tail dependencies in the presence of the unknown marginal distribution of $\varepsilon_{m,t}$ and $\xi_{i,t}$. The threshold τ in equation (A3) is chosen as the unconditional Value-at-Risk of the market equity return, κ in equation (A4) is defined as $\kappa = \frac{VaR_m}{\sigma_{m,t}}$, $K(\cdot)$ is the integral over the Gaussian kernel function $k(\cdot)$ and λ is a bandwidth parameter. For further references with respect to the estimation of the kernel function, we refer to Scaillet (2005).

The unconditional $VaR_m(\psi)$ is in a straightforward manner estimated based on the ex-post realized standardized market equity that is (Benoit et al. (2013), p.43)

$$\widehat{VaR}_m(\psi) = percentile\ \left(\{r_{m,t}\}_{t=1}^{T}, \psi\right) \tag{A7}$$

with ψ equal to 5%.

SRISK

In equation (2.3) in the main text, SRISK is defined as

$$SRISK_{i,t} = k \cdot L_{i,t} - (1-k) \cdot (1 - LRMES_{i,t}) \cdot EQ_{i,t} \tag{A8}$$

with the regulatory capital ratio k, leverage L and the market-valued equity EQ. We do not restrict values at the zero lower bound, as this would result in undistinguishable ranks for banks

with negative SRISK values. The computation of SRISK is straightforward and builds upon MES. Note that several different methods can be used to compute SRISK while we follow Brownlees and Engle (2016). The estimation of the long–run MES is approximated by (see, Acharya et al. (2012), p.61)[25]

$$LRMES_{i,t} \approx 1 - \exp\left(-18 \cdot MES_{i,t}\right). \tag{A9}$$

ΔCoVaR

In equation (2.4) and (2.5) of the main text, CoVaR and ΔCoVaR are defined, respectively, as

$$P\left(r_{i,t} \leq CoVaR_t^{i|r_{m,t}=VaR_{m,t}(\psi)} \middle| r_{m,t} = VaR_{m,t}(\psi)\right) = \psi \tag{A10}$$

with

$$\Delta CoVaR_t(\psi) = -\left(CoVaR_t^{i|r_{m,t}=VaR_{m,t}(\psi)} - CoVaR_t^{i|r_{m,t}=VaR_{m,t}(0.5)}\right). \tag{A11}$$

In our DCC-GARCH framework, this leads to the estimation of ΔCoVaR as (see Benoit et al. 2013), p.44)

$$\Delta\widehat{CoVaR}_t(\psi) = -\hat{\gamma}_{m,t} \cdot \left(\widehat{VaR}_{m,t}(\psi) - \widehat{VaR}_{m,t}(0.5)\right). \tag{A12}$$

Once again, ΔCoVaR is multiplied with (-1) such that larger values represent larger systemic risk. While ψ is set to equal 0.05, an estimate for $\gamma_{m,t}$ can be computed with the help of the previously estimated parameters for the conditional variance and correlation, respectively, that is (see Benoit et al. (2013), p.44)

$$\hat{\gamma}_{m,t} = \frac{\hat{\rho}_{i,t}\hat{\sigma}_{i,t}}{\hat{\sigma}_{m,t}}. \tag{A13}$$

The conditional market $VaR_{m,t}$ is computed with the help of the estimate of the conditional variance $\hat{\sigma}_{m,t}$. Under some general assumptions, Benoit et al. (2013), p.43 show that

$$\widehat{VaR}_{m,t}(\psi) = F_m^{-1}(\psi)\hat{\sigma}_{m,t} \tag{A14}$$

with $F_{m,}$ being the empirical distribution of the (unknown) true distribution of the standardized market returns $\left(\frac{r_{m,t}}{\sigma_{m,t}}\right)$.

[25] Note that Acharya et al. (2012) instead use a threshold of -2% market return.

Technical Appendix B (Essay 2)

Jumps

The construction of the jump-adjusted HAR model strictly follows the approach presented in Andersen et al. (2007). We refer to their work for theoretical underpinnings and to the work of Prokopczuk et al. (2016) for the notation and a detailed explanation of the empirical estimation, as well as the cited references therein.

Loosly speaking, a logarithmic price process is divided into a continuous part and a jump component. The underlying process is modelled as a so-called jump-diffusion process in continuous time, that is (see Andersen et al. (2007), p.702)

$$dp_t = \mu_t dt + \sigma_t dW_t + \eta_t dY_t \tag{B1}$$

with continuous drift variable μ_t, σ_t as a stochastic volatility component, the Brownian motion W_t, Y_t being a count process for jumps that similarly to an indicator variable equals to 1 in case a jump occurs and zero otherwise, and η_t being the size of the jump, i.e. the size of the change of prices. The quadratic variation of the cumulative return process derived from the given prices is eventually estimated with the help of the realized volatility estimator RV. First, the quadratic variation of the given price process can be expressed as the sum of continuous and non-continuous components. That is, (following Andersen et al. (2007), p.702)

$$QV_t = \int_{t-1}^{t} \sigma_s^2 ds + \sum_{t-1<\tau_i\leq t} \eta_{\tau_i}^2. \tag{B2}$$

The first summand of the right-hand side corresponds to the continuous fraction of the quadratic variance, the so called integrated variance. The second summand corresponds to the jump component where τ_t expresses time intervals where jumps occur.

Second, it can be shown (see e.g. Andersen et al. (2003)) that the daily quadratic variation of the underlying price process can be consistently estimated by the realized variance for intraday intervals of infinitesimal length, that is (see, Prokopczuk et al. (2016), p.761)

$$QV_t = \lim_{j\to\infty} RV_t \tag{B3}$$

where j is the number of intraday intervals. Barndorff-Nielsen and Shepard (2004) show that the continuous part of the quadratic variation can be consistently estimated by the so-called bi-power variation (BPV) (following Prokopczuk et al. (2016), p.761)

$$BPV_t = \mu_1^{-2} \left(\frac{J}{J-(k+1)}\right) \sum_{m=k+2}^{J} |r_{t_{m-(k+1)}}| \cdot |r_{t_m}| \tag{B4}$$

where $\mu_1 = E(|Z|) = \sqrt{\frac{2}{\pi}}$ is the mean of the absolute value of a standard normally distributed random variable Z. k is the number of staggered returns because it can be shown that in particular the first intraday return of a day incorporates a lot of microstructre noise. We follow e.g. Prokopczuk et al. (2016), p. 762 and set $k = 1$. Similar to before, it can be shown that for intraday intervals of infinitesimal length, it holds that

$$\lim_{j\to\infty} BPV_t = \int_{t-1}^{t} \sigma_s^2 ds. \tag{B5}$$

Finally, it holds that the difference between the estimators of the quadratic variation (RV) and the integrated variance (BPV) converge in probability to the jump component, that is

$$RV_t - BPV_t \xrightarrow{p} \sum_{t-1<\tau_i\leq t} \eta_{\tau_i}^2. \tag{B6}$$

Based on Andersen et al. (2007) and Huang and Tauchen (2005), we follow Prokopczuk et al. (2016), p. 762 and use the test statistics

$$z_{TQ,t} = \Delta^{-0.5} \frac{(RV_t - BPV_t)/RV_t}{\sqrt{\left(\left(\frac{\pi}{2}\right)^2 + \pi - 5\right) \cdot max\left(1, \frac{TQ_t}{BPV_t^2}\right)}} \tag{B7}$$

to detect significant jumps, where Δ is the length of an intraday interval and TQ_t is the so-called realized tri-power quarticity, defined as (see Prokopczuk et al. (2016), p.762)

$$TQ_t = J\left(\frac{J}{J-2(k+1)}\right) \cdot \mu_{4/3}^{-3} \sum_{m=3}^{J} |r_{t_{m-2(k+1)}}|^{4/3} \cdot |r_{t_{m-(k+1)}}|^{4/3} \cdot |r_{t_m}|^{4/3} \tag{B8}$$

with $\mu_{4/3} \equiv 2^{2/3} \cdot \Gamma(7/6) \cdot \Gamma(0.5)^{-1} = E(|Z|)^{4/3}$. Given the test statistic presented above, the occurrence of jumps is measured as

$$U_t = I_{\{Z_{TQ,t} > \Phi_{1-\psi}\}} \cdot (RV_t - BPV_t) \tag{B9}$$

where $I_{\{Z_{TQ,t} > \Phi_{1-\alpha}\}}$ in an indicator function for the presence of jumps with $\Phi_{1-\psi}$ being the critical value of the cumulative standard normal distribution. Based on Bajgrowicz et al. (2016) we follow Prokopczuk et al. (2016), p.762 and set $\alpha = 0.001$.[26] We refer to the references citied in this chapter for a more in-depth discussion of the derivation and interpretation of the various steps involved in the computation of jumps.

[26] The jump modeling is implemented based on the R-package "highfrequency". See also Boudt et al. (2013) for details regarding the implementation of this package.

Flexible Fourier Form

The estimation of the seasonally normalizing parameter $\hat{s}_{t,j}$ is mainly based upon work from Andersen and Bollerslev (1998), Bollerselv et al. (2000) and a recent application thereof in Behrendt and Schmidt (2018). We refer to their work an the references therein for a more detailed discussion. The main idea is to determine a seasonal trend in a return process via a so-called flexible fourier form (FFF). The latter is estimated via a two-step procedure (see e.g. Hautsch (2012) for a discussion of advantages and shortcomings of the two-step procedure). In order to so, the 10 min-intraday returns are decomposed as follows (see Andersen et al. (1998), p.237):

$$r_{t,j} - E(r_{t,j}) = s_{t,j}\sigma_{t,j}Z_{t,j} \tag{B10}$$

with $E(r_{t,j})$ being the unconditional mean within the chosen sample, $\sigma_{t,j}$ as a daily volatility factor and $Z_{t,j}$ an iid mean zero and unit variance innovation term. Solving equation (B10) for $\hat{s}_{t,j}$, taking squares and applying a logarithmic transformations yields (see Behrendt and Schmidt (2018), p.365)

$$\begin{aligned} ln(\hat{s}_{t,j}^2) = 2ln\frac{|r_{t,j}-\bar{r}|}{\hat{\sigma}_t/\sqrt{J}} = \quad & c + \delta_{0,1}\frac{j}{J_1} + \delta_{0,2}\frac{j^2}{J_2} \\ & + \sum_{p=1}^{P}\left(\delta_{c,p}cos\frac{2\pi p}{J}j + \delta_{s,p}sin\frac{2\pi p}{J}j\right) + \nu_{t,j}. \end{aligned} \tag{B11}$$

The fourier form on the right-hand side of the equation consists of a constant c, two normalizing constants $J_1 = (J+1)/2$ and $J_2 = (J+1)(J+2)/6$, the absolute number of intraday intervals J and a so-called tuning parameter p that determines the order of the fourier expansion. There is no ultimate consensus in the literature regarding the optimal p. For the purpose of this study it is not meaningful to estimate an optimal value for p for each individual time series of stock returns since this will necessarily result in a loss of generalization of our results. Therefore we follow prominent examples in the literature (see e.g. Bollerselv et al. (2000), p.44) and set $p = 6$.[27] The estimate of the daily volatility factor $\hat{\sigma}_{t,j}$ is conveniently obtained via $\hat{\sigma}_{t,j} = \hat{\sigma}_t/\sqrt{J}$. This is also the reason why we use a two-step estimation procedure. First, an estimate of the daily volatility factor is required before in a second step the FFF regression can be estimated

[27] Note that we refrain from incorporating indicator variables for specific macroeconomic announcements as in Andersen and Bollerslev (1998) or Bollerslev et al. (2000) since this is not the focus of our work. We stick to the rather general FFF as also employed in Behrendt and Schmidt (2018).

via plain vanilla OLS. There is no consensus in the literature regarding the specific estimation technique with respect to the daily volatility factor, but often a GARCH-type model variant is used. Once again, we advocate the use of a simple and widely used method that might be well suited to be applied to several different time-series. Therefore, we stick to the classical GARCH(1,1) model that is estimated based on an enlarged sample of daily returns from January 2010 to April 2016, retrieved from Thomson Reuters Datastream.

Finally, the fitted values of the right-hand side of the FFF regression denoted by $\hat{x}_{t,j}$ are used to compute the standardized seasonal component $\hat{s}_{t,j}$ via (see Behrend and Schmidt (2018), p. 366)

$$\hat{s}_{t,j} = TJ \frac{exp(\hat{x}_{t,j}/2)}{\sum_{t=1}^{T} \sum_{j=1}^{J} exp(\hat{x}_{t,j}/2)}. \tag{B12}$$

Once again, we refer to the references citied in this chapter for a more in-depth discussion of the derivation and interpretation of the various steps involved in disentangling a seasonal component from the observed intraday volatility process with the help of a flexible fourier transformation.

Technical Appendix C (Essay 3)

Proof of Proposition 1

$\partial q_u/\partial\alpha = -\frac{(N-1)(hcN+(N-1)\theta)}{(N(1+c-\alpha)+\alpha)^2} < 0$, $\partial q_u/\partial\theta = \frac{1-N}{N(c+1-\alpha)+\alpha} < 0$ and

$\partial q_u/\partial N = -\frac{hc\alpha+\theta+c\theta}{(N(1+c-\alpha)+\alpha)^2} < 0$.

Proof of Proposition 2

The difference $q_o - q_u = \frac{(N-1)(hc\alpha+\theta+c\theta)}{(1+c)(N(1+c-\alpha)+\alpha)} > 0$, $\partial(q_o - q_u)/\partial\theta = \frac{N-1}{N(c+1-\alpha)+\alpha} > 0$ and $\partial(q_o - q_u)/\partial\alpha = \frac{(N-1)(hcN+(N-1)\theta)}{(N(1+c-\alpha)+\alpha)^2} > 0$.

Proof of Proposition 3

The individual effort levels of signatories are given by:

$$q_s^* = \frac{h\Big((n-1)n\alpha - (N(\alpha-1)-n\alpha)\big((1+c)N-(n+N-1)\alpha\big)\Big) - (n-N)(-(1+c)N+(n+N-1)\alpha)\theta}{(1+c-\alpha)((1+c)N^2+(n-N)(n+N-1)\alpha)}, \tag{C1}$$

whereas the individual effort levels of free riders are given by:

$$q_f^* = \frac{-hN^2(1+c-\alpha)(\alpha-1)+hN(1+c-\alpha)\alpha+hn\alpha(n-1+\alpha-n\alpha) - (1+c)(N-1)N\theta+(N-n)(n+N-1)\alpha\theta}{(1+c-\alpha)((1+c)N^2+(n-N)(n+N-1)\alpha)}. \tag{C2}$$

Inserting the equilibrium effort levels into equations (4.12) and (4.15) we get the equilibrium net benefits given by:

$$\pi_s = (h^2((1+c)^2N^4 + (1+c)N^2(2(n-1)n + (2-3N)N)\alpha + (n - N)(n + n^3 + n^2(N-2) + (c^2-3)nN^2 + N(-1 + (4 + c^2(N-2) - 3N)N))\alpha^2 + (c - 1)(n-N)^2(n+N-1)^2\alpha^3) + 2hc(n-N)\alpha((1 + c)N^2(n+N-2) + (n-N)(n+N-1)^2\alpha)\theta + (1 + c)(n-N)((1+c)N^2(n+N-2) + (n - N)(n+N-1)^2\alpha)\theta^2)/(2(1+c - \alpha)((n-1)n\alpha + N(N(1+c-\alpha)+\alpha))^2) \tag{C3}$$

and

$$\pi_f = (h^2(2N^3(1+c-\alpha)\alpha(1+(c-1)\alpha) + (n-1)^2n^2\alpha^2(1+(c - 1)\alpha) + 2(n-1)nN\alpha^2(1+(c-1)\alpha) - N^4(1+c - \alpha)(\alpha-1)(1+c+(c-1)\alpha) + N^2(1+c-\alpha)\alpha(\alpha - c\alpha + 2(n-1)n(1+(c-1)\alpha))) - 2hc\alpha((1 + c)(-2(n-1)n + (N-1)^2)N^2 - (n-N)^2(n+N-1)^2\alpha)\theta - (1+c)((1+c)(-2(n - 1)n + (N-1)^2)N^2 - (n-N)^2(n+N-1)^2\alpha)\theta^2)/(2(1+c - \alpha)((n-1)n\alpha + N(N(1+c-\alpha)+\alpha))^2). \tag{C4}$$

If $n = 1$ effort levels of signatories and free riders are the same and coincide with the non-cooperative Nash equilibrium. Hence, a comparison is only valid if $n \geq 2$. Assuming $n \geq 2$ the following effort level differences are given by:

$$q_s^* - q_c = \frac{(n-1)\big((1+c-\alpha)N^2 + N\alpha - n\alpha\big)(hc\alpha + \theta + c\theta)}{(1+c-\alpha)(N(1+c-\alpha)+\alpha)\big((1+c-\alpha)N^2 + (n^2-n+N)\alpha\big)} > 0, \tag{C5}$$

$$q_f^* - q_c = -\frac{(n-1)n\alpha(hc\alpha + \theta + c\theta)}{(1+c-\alpha)(N(1+c-\alpha)+\alpha)\big((1+c-\alpha)N^2 + (n^2-n+N)\alpha\big)} < 0, \tag{C6}$$

and

$$\bar{q} - q_c = \frac{(n-1)nN(hc\alpha + \theta + c\theta)}{(N(1+c-\alpha)+\alpha)\big((1+c-\alpha)N^2 + (n^2 - n + N)\alpha\big)} > 0. \tag{C7}$$

Hence, $q_s^* > \bar{q} > q_c > q_f^*$ if $n \geq 2$.

$$\partial q_s^* / \partial n = \frac{N(-n^2\alpha + N(N(1+c-\alpha)+\alpha))(hc\alpha + \theta + c\theta)}{(1+c-\alpha)((1+c-\alpha)N^2 + (n^2 - n + N)\alpha)^2}, \tag{C8}$$

is > 0 for $n < \frac{\sqrt{N}\sqrt{N(1+c-\alpha)+\alpha}}{\sqrt{\alpha}}$ and < 0 for $n > \frac{\sqrt{N}\sqrt{N(1+c-\alpha)+\alpha}}{\sqrt{\alpha}}$. Clearly q_s^* is increasing in n if $\frac{\sqrt{N}\sqrt{N(1+c-\alpha)+\alpha}}{\sqrt{\alpha}} > N$. Rewriting this condition, we get $c > -1 + \alpha\left(2 - \frac{1}{N}\right)$. This is fulfilled for $\alpha \leq \frac{1}{2}$ or for $c \geq 1$.

$$\partial q_f^* / \partial n = -\frac{(2n-1)N\alpha(hc\alpha + \theta + c\theta)}{(1+c-\alpha)\big((1+c-\alpha)N^2 + (n^2 - n + N)\alpha\big)^2} < 0. \tag{C9}$$

This means that the effort levels of the free riders strictly decrease in the number of cooperating countries.

$$\partial \pi_s / \partial n = \frac{N^2((n-1)N^2(1+c-\alpha) + n^3\alpha - nN\alpha)(hc\alpha + \theta + c\theta)^2}{(1+c-\alpha)((1+c-\alpha)N^2 + (n^2 - n + N)\alpha)^3} \tag{C10}$$

is definitely > 0 for $n > \sqrt{N}$ and < 0 for $n < \sqrt{N}$.

$$\partial \pi_f / \partial n = \frac{(2n-1)N^2(N^2(1+c-\alpha)+\alpha)(hc\alpha + \theta + c\theta)^2}{(1+c-\alpha)((1+c-\alpha)N^2 + (n^2 - n + N)\alpha)^3} > 0. \tag{C11}$$

Assuming again $n \geq 2$ the difference of net benefits between free riders and signatories is given by:

$$\pi_f - \pi_s = \frac{(n^2-1)N^2(hc\alpha + \theta + c\theta)^2}{2(1+c-\alpha)\big((1+c-\alpha)N^2 + (n^2 - n + N)\alpha\big)^2} > 0. \tag{C12}$$

Proof of Proposition 4

Inserting equations (C3) and (C4) into the internal stability function $\pi_s(n) - \pi_f(n-1) \geq 0$ we get:

$$\frac{\begin{array}{c}-(((n-1)N^2((n-3)N^4(1+c-\alpha)^2 + 2(n-1)N^3(1+c-\alpha)\alpha \\ +(n-3)(n-1)^2n^2\alpha^2 + 2n(-1+(n-2)n)N\alpha^2 \\ +N^2\alpha(2(1+c)(-2+(n-3)(n-1)n) \\ +(5+n(-5-2(n-4)n))\alpha))(hc\alpha+\theta+c\theta)^2)\end{array}}{(2(1+c-\alpha)((1+c)N^2+(n-N-1)(n+N-2)\alpha)^2((1+c-\alpha)N^2+(n^2-n+N)\alpha)^2))} \geq 0 \tag{C13}$$

Inserting equations (C3) and (C4) into the external stability function $\pi_f(n) - \pi_s(n+1) \geq 0$ we get:

$$\frac{\begin{array}{c}(nN^2((1+c)^2(n-2)N^4 - 2(1+c)N^2(2-(n-2)n(1+n) - nN + (n-2)N^2)\alpha \\ +(1+n-N)(-4N+(n+N)((n-2)n(1+n) \\ +(2+n)N-(n-2)N^2))\alpha^2)(hc\alpha+\theta+c\theta)^2)\end{array}}{(2(1+c-\alpha)((1+c-\alpha)N^2+(n^2-n+N)\alpha)^2(n(1+n)\alpha+N(N(1+c-\alpha)+\alpha))^2))} \geq 0. \tag{C14}$$

substituting $N = 2$ and $n = 2$ into equations (C13) and (C14) gives:

$$\frac{(hc\alpha+\theta+c\theta)^2}{2(1+c)(-2-2c+\alpha)^2} > 0 \text{ and } \frac{\alpha(2+2c+3\alpha)(hc\alpha+\theta+c\theta)^2}{8(1+c)^2(1+c-\alpha)(1+c+\alpha)^2} > 0. \tag{C15}$$

Clearly, both conditions are fulfilled and, hence, two countries cooperate.

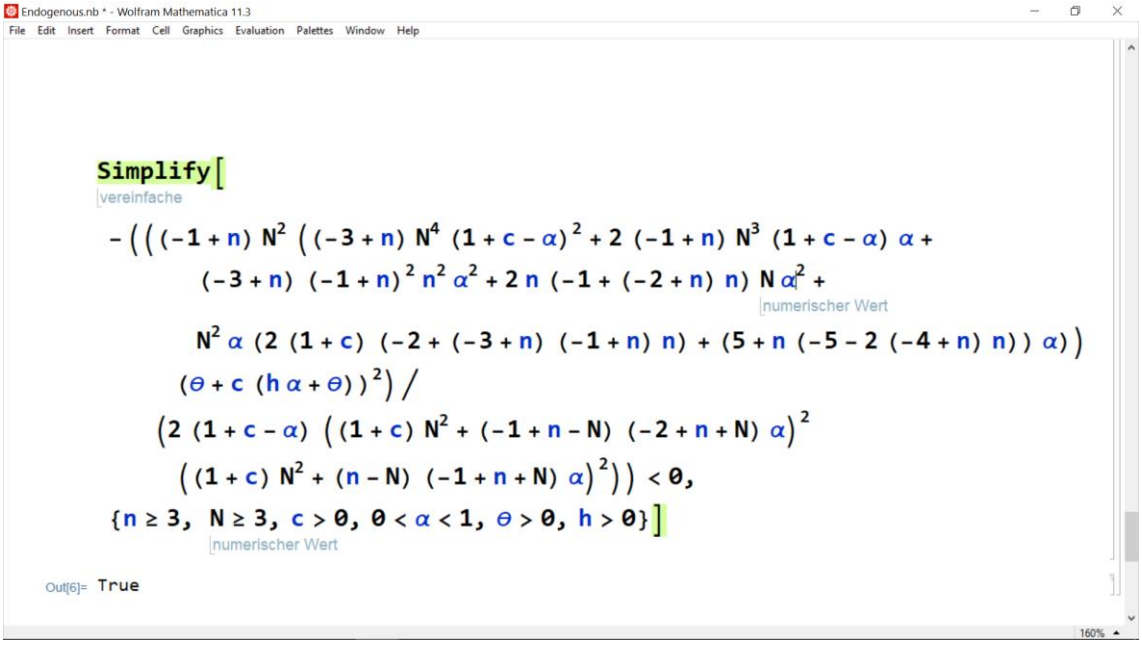
Then we analyze the stability of a two-country coalition if $N > 2$. Substituting $n = 2$ into equation (C13) gives:

$$\frac{\left((1+c)^2N^4 - 2(1+c)N^2(N+N^2-4)\alpha + (N-2)(N-1)\left(2+N(5+N)\right)\alpha^2\right)(hc\alpha+\theta+c\theta)^2}{2(1+c-\alpha)(N(1+c-\alpha)+\alpha)^2\left(2\alpha+N(N(1+c-\alpha)+\alpha)\right)^2} \tag{C16}$$

Internal stability is fulfilled if $c \geq -1 + \frac{(N^2+N-4)\alpha}{N^2} + \frac{2\alpha\sqrt{3+(N-3)N}}{N^2}$. We note that the right side of the inequality increases in α if $N \geq 4$ and decreases in N if $N \geq 5$. Substituting $n = 2$ into equation (C14) gives:

$$\frac{4N^3\alpha((1+c)(N^2-N) - (N^2-2N-3)\alpha)(hc\alpha+\theta+c\theta)^2}{(1+c-\alpha)(2\alpha+N(N(1+c-\alpha)+\alpha))^2(6\alpha+N(N(1+c-\alpha)+\alpha))^2} > 0, \tag{C17}$$

which implies that external stability of $n = 2$ is always fulfilled. This implies that a three-country coalition is not internally stable. We further show that coalitions of size $n \geq 3$ are not internally stable. Numerical proof with Mathematica software package:



The last line contains parameter values to be substituted in the inequality above. The result indicates "True", that is the internal stability for the given parameterization is never fulfilled. Hence stable coalitions of size $n \geq 3$ do not exist. If $N \geq 3$ either $n = 2$ or $n = 1$ is stable in equilibrium. From that we note that incentives to cooperate do not increase in α and θ.

Aggregate payoffs are $\Pi_a = n\pi_s + (N - n)\pi_f$. By inserting $n = 0$ into (C4) we get the aggregate net benefit of the non-cooperative case:

$$\Pi_u = -\frac{N\begin{pmatrix}(-hN + h(N-1)\alpha + (N-1)\theta)(h(1+c)N + \\ h(c-1)(N-1)\alpha + (1+c)(N-1)\theta)\end{pmatrix}}{2(N(1+c-\alpha)+\alpha)^2}, \tag{C18}$$

and by inserting $n = N$ in (C3) we get the net benefit of the social optimal case:

$$\Pi_o = \frac{Nh^2}{2 + 2c}. \tag{C19}$$

Inserting the net benefits of the different scenarios into (4.21) and assuming $n = 2$ we obtain:

$$CGI = \frac{\begin{pmatrix} 4(1+c)N^4(1+c-\alpha)^2 - 4(1+c)N^2(1+c-\alpha)\alpha + 4(1+c)\alpha^2 \\ -2(1+c)N\alpha^2 + 2(1+c)N^3(-3(1+c)^2 + 8(1+c)\alpha - 5\alpha^2) \end{pmatrix}}{\begin{pmatrix} H^6(1+c-\alpha)^3 - 2H^5(1+c-2\alpha)(1+c-\alpha)^2 + H^4(1+c-\alpha)((1+c-\alpha)^2+\alpha^2) \\ -2H^3(3+3c-4\alpha)(1+c-\alpha)\alpha + H^2(4+4c-7\alpha)(1+c-\alpha)\alpha - 4H(1+c-\alpha)\alpha^2 \end{pmatrix}}. \tag{C20}$$

We note from equation (C20) that the CGI is invariant in h and θ and that $\lim_{N\to\infty} CGI = 0$ if the number of countries interconnected is very large.

Proof of Proposition 5

Inserting equations (C3) and (C4) into the altered internal stability function $\pi_s(n) + g(n-1) - \pi_f(n-1) \geq 0$ and substituting $n = N$ we get:

$$\frac{1}{2}(N-1)\left(2g - \frac{\big((1+c)(N-3)N^2 + 4(N-1)\alpha\big)(hc\alpha + \theta + c\theta)^2}{(1+c)(1+c-\alpha)(N(N+cN-2\alpha)+2\alpha)^2}\right) \geq 0. \tag{C21}$$

We note that equation (C22) is strictly increasing in g and hence internal stability is fulfilled if $g \geq \frac{((1+c)(N-3)N^2+4(N-1)\alpha)(hc\alpha+\theta+c\theta)^2}{2(1+c)(1+c-\alpha)(N(N+cN-2\alpha)+2\alpha)^2} \equiv A$. If the club benefits are large enough, the grand coalition can be stable. By multiplying the condition for g by $(N-1)$ we get the amount of club benefits $E_s(N) = g(N-1)$ which are needed to stabilize the grand coalition. The condition then becomes $E_s(N) \geq A(N-1)$. Putting the necessary club benefits in relation with the public benefits given by equation (4.23) we get:

$$\gamma \equiv \frac{E_s(N)}{H_s(N)} \geq \frac{(1+c)(N-1)((1+c)(N-3)N^2 + 4(N-1)\alpha)(hc\alpha + \theta + c\theta)^2}{(1+2c)(1+c-\alpha)\alpha(h(1+c)N^2 - 2h(N-1)\alpha)^2}. \tag{C22}$$

$$\partial\gamma/\partial\theta = \frac{2(1+c)^2(N-1)((1+c)(N-3)N^2 + 4(N-1)\alpha)(hc\alpha + \theta + c\theta)}{(1+2c)(1+c-\alpha)\alpha(h(1+c)N^2 - 2h(N-1)\alpha)^2} > 0$$

and

$$\partial\gamma/\partial h = -\frac{2(1+c)^2(N-1)((1+c)(N-3)N^2 + 4(N-1)\alpha)\theta(hc\alpha + \theta + c\theta)}{h^3(1+2c)(1+c-\alpha)\alpha(N(N+cN-2\alpha)+2\alpha)^2} > 0.$$

Proof of Lemma 1

The maximization problem of the free riders is given by:

$$\max_{q_f} \pi_f = \alpha b\bar{q} + (1-\alpha)bq_f - \frac{1}{2}c{q_f}^2 - \theta\left(q_f - \bar{q}\right). \tag{C23}$$

The corresponding reaction function is given by:

$$q_f^* = \frac{b(N+\alpha-N\alpha)+\theta-N\theta}{cN} \tag{C24}$$

The maximization problem of the signatories is given by:

$$\max_{q_s} \pi_s = \alpha b\left(\frac{nq_s+(N-n)q_f}{N}\right) + (1-\alpha)bq_s + g(n-1) - \frac{1}{2}c{q_s}^2 \tag{C25}$$

$$-\theta\left(q_s - \frac{nq_s+(N-n)q_f}{N}\right)$$

.

The corresponding reaction function is given by:

$$q_s^* = \frac{b(N+n\alpha-N\alpha)+(n-N)\theta}{cN}. \tag{C26}$$

$\partial q_s^*/\partial n > 0$ and $\partial q_f^*/\partial n = 0$ and, hence, $q_s^* > \bar{q} > q_f^* = q_c$. Inserting equilibrium effort levels given by equations (C24) and (C26) into equations (C23) and (C25), we get the equilibrium net benefits given by:

$$\pi_s = \frac{b^2(N^2+(n-N)(N+n-2)\alpha^2) + 2b(n-N)(N+n-2)\alpha\theta + (n-N)(N+n-2)\theta^2}{2cN^2} \tag{C27}$$

and

$$\pi_f = \frac{b^2(N^2-(-2(n-1)n+(N-1)^2)\alpha^2) - 2b(-2(n-1)n + (N-1)^2)\alpha\theta - (-2(n-1)n+(N-1)^2)\theta^2}{2cN^2}. \tag{C28}$$

$\partial\pi_s/\partial n = \frac{(n-1)(b\alpha+\theta)^2}{cN^2}$ and $\partial\pi_f/\partial n = \frac{(2n-1)(b\alpha+\theta)^2}{cN^2} > 0$ and $\pi_f > \bar{\pi} > \pi_s > \pi_c$.

Proof of Proposition 6

Substituting the net benefits given by equations (C27) and (C28) into equation (4.20) we receive the relevant stability conditions. First, we solve the internal stability condition $\pi_s(n) \geq \pi_f(n-1)$ for n, which yields that internal stability is fulfilled for:

$$3 + \frac{2cgN^2}{(b\alpha+\theta)^2} \geq n \geq 1. \tag{C29}$$

Then, we solve the external stability condition $\pi_f(n) \geq \pi_s(n+1)$ for n, which yields that the external stability is fulfilled for:

$$n \geq 2 + \frac{2cgN^2}{(b\alpha+\theta)^2}. \tag{C30}$$

Hence, both conditions are fulfilled simultaneously for:

$$3 + \frac{2cgN^2}{(b\alpha+\theta)^2} \geq n \geq 2 + \frac{2cgN^2}{(b\alpha+\theta)^2}. \tag{C31}$$

If $g = 0$, coalitions of size $n \in \{2,3\}$ are stable. We note from equation (C31) that $\partial n/\partial c$ and $\partial n/\partial N > 0$ and that $\partial n/\partial\alpha$ and $\partial n/\partial\theta < 0$. Both sides of the interval grow proportionally in all exogenous parameters. Hence, it is sufficient to analyze the internal stability to determine the stable size of a coalition. Solving the internal stability with respect to g we get:

$$g \geq \frac{(n-3)(b\alpha+\theta)^2}{2cN^2}. \tag{C32}$$

This equation determines the necessary size of g in order to stabilize a coalition of size n. If we substitute N into (C32) we get the value of g in order to stabilize the grand coalition. By multiplying this by $(N-1)$ we get the amount of club benefits $E_s(N)$ which are needed to stabilize the grand coalition. Putting the necessary club benefits in relation with the public benefits $H_s(N)$ we get:

$$\gamma \equiv \frac{E_s(N)}{H_s(N)} = \frac{(N-3)(N-1)(b\alpha+\theta)^2}{2b^2N^2\alpha}. \tag{C33}$$

$\partial\gamma/\partial\theta = \frac{(N-3)(N-1)(b\alpha+\theta)}{b^2N^2\alpha} > 0$, $\partial\gamma/\partial N = \frac{(2N-3)(b\alpha+\theta)^2}{b^2N^3\alpha} > 0$ and

$\partial\gamma/\partial b = -\frac{(N-3)(N-1)\theta(b\alpha+\theta)}{b^3N^2\alpha} < 0$. Aggregate net benefits are given by:

$$\Pi_a = \frac{\begin{array}{c} b^2(N^3-(N-n)(-n^2+(N-1)^2+nN)\alpha^2) - \\ 2b(N-n)(-n^2+(N-1)^2+nN)\alpha\theta \\ -(N-n)(-n^2+(N-1)^2+nN)\theta^2 \end{array}}{2cN^2}. \tag{C34}$$

By inserting $n = 0$ into (C34) we get the aggregate net benefit of the non-cooperative case:

$$\Pi_u = -\frac{(-bN+b(N-1)\alpha+(N-1)\theta)(b(N+(N-1)\alpha)+(N-1)\theta)}{2cN}, \tag{C35}$$

and by inserting $n = N$ in (C34) we get the net benefit of the social optimal case:

$$\Pi_o = N\left(\frac{b^2}{2c} + g(N-1)\right). \tag{C36}$$

Inserting equations (C34), (C35) and (C36) into equation (4.21) we obtain the CGI given by:

$$CGI = \frac{(n-1)n(2N-n-1)(b\alpha+\theta)^2}{(N-1)N(2cgN^2+(N-1)(b\alpha+\theta)^2)}. \tag{C37}$$

If $g = 0$ we substitute $n = 2$ and $n = 3$ as stable solutions into equation (C37) and get:

$$CGI(2) = \frac{4N-6}{N(N-1)^2} \quad \text{and} \quad CGI(3) = \frac{12(N-2)}{N(N-1)^2}. \tag{C38}$$

We note that equation (C38) does not depend on b, α, c and θ. If $g > 0$ we note from equation (C37) that the CGI depends on all parameters and that the CGI always converges to zero if $N \to \infty$.