UNIVERSITÁ CATTOLICA DEL SACRO CUORE

Sede di Brescia

Interfaculty of

Mathematical, Physical and Natural Sciences,

Banking, Finance and Insurance Sciences

Master of Science in Applied Data Science for Banking and Finance



Modeling systemic risk with Δ-CoVaR: a comparison of linear and neural network approaches

Supervisor:

Prof. Andrea Roncella

Co-Supervisor:

Prof.ssa Velia Cenciarelli

Dissertation by:

Cecilia Paolini

Student ID n.: 5308096

Academic Year 2024/2025

Contents

List of Figures **4**

1 Introduction **7**

2 State of the Art **11**

2.1 Systemic Risk 11

2.2 CoVaR 14

2.3 Machine Learning for CoVaR 15

2.4 Systemic Risk Network Techinques 16

3 Dataset **20**

3.1 Data Selection 20

3.1.1 BankFocus 20

3.1.2 Bloomberg 23

3.2 Market Variables Calculations 26

3.3 Panel Regression Variables 30

4 Methodology **34**

4.1 ΔCoVaR overview 34

4.2 Implementation 40

4.2.1 ΔCoVaR Models 44

4.2.2 Forward-ΔCoVaR 49

5 Performance Comparison **53**

5.1 Pinball Loss Function 53

5.2 Negative Residuals 56

5.3 Pseudo R-Squared 59

5.4 Conclusion 62

6 Graphical Analysis **64**

6.1 Entire period 64

6.2 Crisis Periods 70

7 Forward-ΔCoVaR **80**

7.1 Correlation Matrix 80

7.2 Panel regression 84

7.2.1 Comparisons 85

7.2.2 Coefficients 86

7.3 Conclusion 90

8 Final Remarks **93**

Bibliography **97**

List of Figures

5.1 Pinball loss frequency histogram of 5%VaR with LQR (Active) 54

5.2 Pinball loss frequency histogram of 5%CoVaR with LQR (Active) 55

5.3 Pinball loss frequency histogram of 5%VaR with NNQR (Active) 55

5.4 Pinball loss frequency histogram of 5%CoVaR with NNQR (Active) 56

5.5 Negative residuals percentage of 5%VaR with LQR (Active) . 57

5.6 Negative residuals percentage of 5%CoVaR with LQR (Active) 58

5.7 Negative residuals percentage of 5%VaR with NNQR (Active) 58

5.8 Negative residuals percentage of 5%CoVaR with NNQR (Active) 59

5.9 Pseudo R-squared of 5%VaR with LQR (Active) 60

5.10 Pseudo R-squared of 5%CoVaR with LQR (Active) 61

5.11 Pseudo R-squared of 5%VaR with NNQR (Active) 61

5.12 Pseudo R-squared 5%CoVaR with NNQR (Active) 62

6.1 Var, CoVaR, ΔCoVaR for commercial banks - US area (Active) 66

6.2 Var, CoVaR, ΔCoVaR for commercial banks - Euro area (Active) 67

6.3 Var, CoVaR, ΔCoVaR for bank holding companies - US area (Rest) 68

6.4 Var, CoVaR, ΔCoVaR for bank holding companies - US area (Inactive) 69
6.5 Var, CoVaR, ΔCoVaR (LQR) for commercial banks - Euro area (Merged) 72
6.6 Var, CoVaR, ΔCoVaR (NNQR) for commercial banks - Euro area (Merged) 73
6.7 Var, CoVaR, ΔCoVaR (LQR) for bank holding companies - US area (Merged) 74
6.8 Var, CoVaR, ΔCoVaR (NNQR) for bank holding companies - US area (Merged) 75
6.9 Var, CoVaR, ΔCoVaR (LQR) for bank holding companies - Euro area (Active) 76
6.10 Var, CoVaR, ΔCoVaR (NNQR) for bank holding companies - Euro area (Active) 77

7.1 Panel regression variables correlation matrix (Active) 81
7.2 Panel regression variables correlation matrix (Merged) 82

Acknowledgements

I would like to express my deepest gratitude to my thesis supervisor, professor Andrea Roncella, for his support throughout every step of my research and for his flexibility with the eight-hour time difference.

I also want to thank Songia Eugenio for patiently sending me countless data excels, which were essential to this work.

A heartfelt thank you to my family, whose constant support made this achievement possible.

And last but not least, a very special thank you to Lachlan, for always brightening up my days and for being my number one fan.

Chapter 1

Introduction

The literature on systemic risk had been present for many years prior to the Great Financial Crisis of 2007-2008 (GFC), but these studies mainly focused on individual risk factors that affect individual financial institutions separately.

Throughout the years following the GFC, the interest of researchers in the topic of systemic risk peaked substantially as the data from the crisis provided great academic groundwork for testing and analyzing new theories. Although scholars are still struggling to find consensus on a common definitive definition of systemic risk, a general agreement point seemed to be that "while credit risk, liquidity risk, operational risk, etc. can be directly attributed to a given institution, systemic risk can only be attributed indirectly" (Smaga, 2014, p. 2) [1]. In fact, post-crisis papers elaborate on the concept that systemic risk cannot be seen as an isolated event, a consequence of financial stress in an institution, it needs to be addressed as a phenom of financial interconnectedness that could start with a single event, for example, an initial shock, which then propagates on a regional or global scale. Addressing interconnection in systemic risk prediction adds a further layer of

complexity to the problem since financial institutions are very often linked to one another, but those links are not always easy to detect in advance. Given the fact that direct connections among financial institutions rely on confidential data, generally very difficult to obtain, researchers developed measures and techniques that circumvent the need for direct exposure information by using market data (e.g. returns or state variables).

Adrian and Brunnermeier (2011) [2] introduced CoVaR, a measure of systemic risk that addresses the key limitations of previous methods. Instead of assessing a single institution's risk, they quantify the additional risk an institution contributes to the entire financial system during times of stress. They understood that systemic risk can arise from either a single "too big-to-fail" institution or a group of smaller institutions exhibiting similar behaviors, which can create a lack of diversification. They also emphasized that risk measures should be forward-looking, capable of detecting the build-up of financial issues that might materialize often magnified during a crisis. Their proposed measure, CoVaR (Conditional Value at Risk), is the Value at Risk (VaR) of the financial system conditional on a specific institution being in distress. The difference between CoVaR in distress and CoVaR in a "normal" or median state is ΔCoVaR, which isolates the risk contribution of the distressed institution. The authors calculated both CoVaRs using Quantile Regression (QR), a method that differs from traditional mean-based regression by modeling how predictors influence a response variable throughout its entire conditional distribution and not just at its average.

Adrian and Brunnermeier's approach to quantile regression is linear and bivariate. Taylor (2000) [3] had previously already introduced a quantile regression method integrated with neural networks, a methodology known as NNQR. However, NNQR has started to become popular only with the rise

of Big Data and powerful GPUs, offering significant advantages. NNQR can map complex, nonlinear relationships between variables and their quantiles without requiring a predefined functional form. It can also identify intricate interactions that standard QR, which assumes linearity or requires manual specification of nonlinear terms, might miss, especially in large and complex datasets.

Although the new literature on applying Machine Learning techniques to systemic risk measures such as ΔCoVaR is available and prolific, a notable gap remains in the methodology. Many studies introduce different approaches to the problem of predicting systemic risk, but it seems that the authors of these studies either substantially alter the metrics and the underlying variables, or choose a different systemic risk measure altogether. This factor poses a challenge in trying to isolate specific performance improvements of the original ΔCoVaR methodology of Adrian and Brunnermeier, which still appears to be the base reference for systemic risk predictions. Additionally, and perhaps most importantly, the new advanced models proposed for systemic risk analysis and prediction show a high degree of complexity, often resulting in a "black-box" problem, with good predictive power but unclear internal mechanisms. This lack of interpretability hinders the ability to derive meaningful conclusions about, for example, why a particular entity is a source of systemic risk. This is a critical limitation for policymakers and regulators, who also require interpretable insights, not only a risk score.

In light of the shortcomings of the current state of the art, the objectives of this research are the following. First, this work provides a comprehensive review of the ΔCoVaR model computed using the original linear quantile regression (LQR), with a specific focus on conducting a more detailed financial analysis of its results. Second, a direct comparison of LQR ΔCoVaR results

and analysis is performed using the same ΔCoVaR model reconstructed using neural matrix quantile regression (NNQR).
By holding the base model and the data constant, this research uniquely isolates and evaluates the impact of the neural network's ability to capture nonlinear dependencies in the systemic risk model of Adrian and Brunnermeier. Furthermore, the focus on interpreting and providing financial meaning to the results gives additional insight into the predictive power of ΔCoVaR.

The remaining chapters of this thesis are organized as follows. Chapter 2 provides an overview of the most relevant literature on systemic risk, from its definition to the evolution of systemic risk measures after Adrian and Brunnermeier. Chapter 3 describes the data used for all the calculations and models used in the study. Chapter 4 is an extensive description of the methodologies adopted, from the base model of ΔCoVaR to its implementation and the neural network variation employed, ending with a description of the panel regression for forward-ΔCoVaR. Chapter 5 and Chapter 6 provide, respectively, a quantitative performance comparison and a more graphical analysis to evaluate the difference between LQR ΔCoVaR and NNQR ΔCoVaR. Chapter 7 describes the analysis of the results of forward-ΔCoVaR, which shows how the use of different versions of quantile regression influences ΔCoVaR predictions. Finally, Chapter 8 outlines the final thoughts on the work, identifies possible shortcomings in this study, and suggests potential improvements to be made in future research.

Chapter 2

State of the Art

2.1 Systemic Risk

According to Allen and Carletti (2013) [4], before the great financial crisis that began in the summer of 2007, the general idea in the financial world was that preventing each bank from taking large risks would also prevent a build-up risk in the entire financial system. This view turned out to be flawed as it completely ignores systemic risk. As mentioned in the Introduction, the exact definition of systemic risk is still a debating topic in financial studies. Allen and Carletti view systemic risk as a combination of these four factors:

1. **Panics**

   Historically, panics have been a significant source of systemic risk. In the "normal" state, all depositors believe that the other depositors withdraw their funds only according to their consumption needs, which leads to a good equilibrium: banks can satisfy all depositors' demands without liquidating any of the long-term assets. However, if depositors believe that other depositors will withdraw prematurely, then all depositors find it rational to redeem their claims and a panic occurs.

This mechanism activates a snowballing chain of financial events that exacerbates the financial crisis.

2. **Banking crises due to asset price**

   The authors list several reasons for the fall of assets held by banks:

   - Outgrow of the business cycle: if depositors receive information about the impending downturn in the cycle, they will anticipate difficulties in the banking sector and try to withdraw their funds prematurely, precipitating the crisis.
   - Bursting of real estate bubbles: many financial institutions experience severe problems after the collapse in the securitized mortgage market, and these financial issues then easily spread to the real economy.
   - Mispricing due to inefficient liquidity provision and limits to arbitrage: mispricing of assets can generate even for the prices of safe assets to fall below their fundamental values.
   - Sovereign default: for example, a country's financial default like Greece can put pressure on the European financial system, which then puts pressure on the Euro
   - Increases in interest rates: when interest rates start to rise, as a result of policy moves by central banks to restrain inflation or market moves in anticipation of inflation, then the price of all securities, including government debt, will fall, often causing significant solvency problems for banks.

3. **Contagion**

   This is the possibility that distress from one financial institution will spread to others, potentially leading to a system-wide crisis. Allen and

Carletti (2013) [4] mention the failure of Lehman Brothers in 2008 as a prime example, as it led to a loss of confidence in many financial firms and a significant spillover into the real economy. Contagion can also arise from common asset exposures, as private diversification can increase systemic risk by making portfolios more similar. Although greater connectivity in financial networks can reduce the likelihood of widespread default, shocks can have a much larger impact when they do occur.

4. **Foreign exchange mismatches in the banking system**

   Before the 1997 Asian crisis, for example, many banks and firms had borrowed in foreign currencies and were then unable to borrow them when the crisis hit. A lack of foreign exchange reserves prevented central banks from helping, forcing countries like South Korea to turn to the IMF, resulting in a damaging recession.

Smaga (2014) [1] defines systemic risk as the risk that a shock will lead to a significant materialization of imbalances that spread and affect the financial system, ultimately affecting the real economy. The author also distinguishes *systemic risk* from *systematic risk*. According to Smaga, systematic risk is regarded as speculative risk that affects only changes in the prices of financial instruments, while systemic risk is an indirect result of the aggregation and correlation of all individual banks. The author also emphasizes that contagion is a key mechanism for the materialization of systemic risk and can spread through various channels, including bilateral exposures, market factors, and international links.

2.2 CoVaR

Adrian and Brunnermeier (2011) [2] are the first to propose a measure of systemic risk that addresses the main issues in systemic risk modeling. Earlier metrics focused on calculating the risk of individual institutions or mapping the exact path of propagation through often hidden links. Adrian and Brunnermeier (2011), on the other hand, focused on quantifying how much risk a single institution adds to the entire system when it experiences stress. In their article, a key focus of this research, they state that a group of 100 institutions that act like clones can be as precarious and dangerous to the system as the large merged identity. According to the authors, this means that the possible failure of a single 'too big to fail' institution is not the only reason that can lead to spread of risk. A combination of small undetected institutions employing similar strategies and their consequent lack of diversification can also lead to substantial financial risk on a large scale. The authors also add that risk measures should recognize that risk typically builds up in the background in the form of imbalances and bubbles and materializes only during a crisis. This means that measures should not focus only on current and easy observable market movements but should also be able to detect in advance dangerous build-up that might materialize only during a crisis.
Following this rationale, Adrian and Brunnermeier take a common risk measure, the value at risk (VaR) and add the prefix 'Co' which can stand for conditional, contagion, or comovement. According to Jorion (2007) [5] and Kupiec (2002) [6], VaR measures maximum loss in monetary units within a q confidence interval. The CoVaR of the institution i relative to the system is defined, in Adrian and Brunnermeier's words, as the VaR of the whole financial sector conditional on institution i being in distress. ΔCoVaR is the difference between CoVaR conditional on i being in distress and CoVaR

of i being in a 'normal' state, which captures the additional risk that the institution i in distress inflicts on the financial system.

In their paper, Adrian and Brunnermeier calculate CoVaR using quantile regression (QR), a statistical methodology introduced by Koenker and Bassett (1978) [7] that extends traditional mean regression, typically estimated using ordinary least squares (OLS). Traditional regression methods focus solely on modeling the conditional mean of the response, while QR allows investigation of how the predictors influence the response variable across its entire conditional distribution, examining effects at, for instance, the median (0.5 quantile), lower quantiles (e.g., 0.1, 0.25) or upper quantiles (e.g., 0.75, 0.9).

2.3 Machine Learning for CoVaR

Adrian and Brunnermeier's (2011) [2] original CoVaR approach is, however, limited to a linear bivariate context. Taylor (2000) [3] wrote the first widely recognized article that integrated the two methodologies of neural networks (NN) and quantile regression together (NNQR). However, it was not until the mid-2000s, driven by the availability of massive datasets (Big Data), the power parallel processing offered by Graphic Processing Units (GPUs), and algorithmic improvements, that neural networks and derivative methods exploded in popularity in finance and economics.

According to Taylor (2000) [3] and Cannon (2011) [8] one of the main advantages of computing quantile regression through NN is its ability to map complex and nonlinear relationships between input features and the desired quantiles, without specifying the exact functional form in advance. Another advantage of NN is that it is able to understand complex interactions (e.g. $X_1 * X_2$ in regular regression) without specified inputs. Regular QR, on the other hand, assumes either a linear relationship or manual input of nonlin-

ear terms (like polynomials) known beforehand, and explicit specification of any interactions. In a large data set with many variables, regular QR could not be accurate and, if the model is oversimplified, could lead to misleading results.

The study by Keilbar and Wang (2019) [9] introduces a neural network quantitative regression model to address non-linearity in systemic risk. However, their work deviates substantially from an Adrian and Brunnermeier CoVaR model, making the two methodologies difficult to compare. Keilbar and Wang, in fact, use new network-based metrics such as the systemic network risk index (SNRI), the systemic fragility index (SFI), and the systemic risk index (SHI). Another metodology difficult to compare to Adrian and Brunnermeier's is that of Kumar et al. (2024) [10]. Kumar et al. reframed systemic risk analysis by changing the original continuous prediction problem into a classification problem. Rather than providing a continuous score, they categorize banks into discrete classes (high vs. moderate systemic risk emitters). The study results also show a consistent lack of a single set of "most important" variables, since the feature rankings differ significantly between the various machine learning models (LR, RF, GBM, and NN). The Neural Network model of Kumar et al. in particular performed poorly, showing the lowest Area Under the Curve (AUC) in most settings, suggesting it is not a suitable technique for their specific dataset.

2.4 Systemic Risk Network Techinques

An interesting line of systemic risk analysis that requires more modern and sophisticated techniques focuses on estimating the risk spill over the network. Wu et al. (2020) [11] analyze interconnectedness, systemic risk, and influencing factors within the financial system. The authors create a dynamic

network based on stock price data from publicly listed financial institutions. They use Principal Component Analysis (PCA) to measure general connectivity between financial institutions. According to Wu et al., if a large proportion of total return fluctuations are being explained by a small number of principal components, it indicates a high degree of interconnectedness and, thus, a higher potential for systemic risk. Therefore, the nature of systemic risk is determined by observing how the explained variance of the principal components changes over time. To also determine the direction and degree of risk contagion, the study uses linear and nonlinear Granger causal tests. This test measures the causal relationships between the time series of each institution's stock returns, building a directed network where a link from institution j to i exists if j Granger-causes i. Wu et al., to identify key institutions, deconstruct this network to analyze its overall characteristics, such as density and average path length, as well as individual node characteristics like various forms of centrality. Finally, they used an econometric model to investigate how micro- and macroeconomic factors influence the network's connectedness.

The first main issue with Wu et al.'s methodology is the reliance on market returns. Adrian and Brunnermeier argued in their paper that the ultimate concern of systemic risk is directly related to the supply of credit to the real economy. According to Adrian and Brunnermeier using total assets instead of market returns is therefore a better option. The second issue with Wu et al.'s study is the Granger causality approach. This approach, while powerful in identifying causal relationships, can be very complex and may not directly quantify the marginal contribution of a single institution to total systemic risk in the same way as CoVaR and ΔCoVaR. The third problem with Wu et al.'s work is that the authors analyze the factors that influence connect-

edness using an econometric model. However, they do not propose a single predictive "forward" metric: they focus only on studying the relationship between current firm characteristics and contemporaneous connectedness. Therefore, it is not clear in the article by Wu et al. how their study fits into the predicting risk framework.

Graph Neural Networks (GNNs) are another method proposed by Balmaseda et al. (2023) [12] to analyze the systemic risk network. GNNs are a class of deep learning algorithms designed to handle structured data as graphs, which makes them very well suited for analyzing financial networks. While in traditional machine learning models each observation is treated independently, GNNs take advantage of all information available, including node interrelations and node, edge, and graph features. With this method, the graph itself models a systemic risk metric that captures complex and non-linear interactions between the financial entities in the sample. Balmaseda et al. use GNNs to perform a node classification task. They first run a contagion algorithm called DebtRank, which estimates the additional system stress caused by the default of each entity, and then orders the entities by their additional system stress and classifies them into quartiles from 0 to 3, where 0 is the most important class and 3 is the least. The model then learns from a simple discrete classification to predict a continuous risk score (quantile) for each node.

Balmaseda et al.'s methodology seems promising in theory, as it is based on the interconnectedness of entities, which is the core of systemic risk predictions and although the results show a significant improvement in performance over traditional machine learning models. However, the technique has significant limitations. The most obvious one is the pre-existing need for network data. As mentioned above, bilateral exposure data, a crucial ele-

ment of financial networks, are confidential and often not publicly available. The authors address the issue using network reconstruction methods, adding another layer of modeling challenges and another source of potential error. The other issue with GNNs is the need for an extensive data sample, since, according to the authors, the model sometimes critically underperforms and struggles with generalization. The main causes of this issue states by Balmaseda et al are: limited data, susceptibility to hyperparameter tuning, which characterizes all deep learning models as well, and computational requirements, because running large networks requires a powerful GPU which is not accessible to all researchers and small institutions.

Chapter 3

Dataset

Data collection posed many challenges due to the large number of entities required for analysis and the limited remote access to major databases for university students. For this analysis, with the help of the university library, a combination of the **BankFocus** and **Bloomberg** databases were used.

3.1 Data Selection

3.1.1 BankFocus

The selection of relevant entities and relevant characteristics was first done on BankFocus. Moody's Analytics BankFocus is a database of banks worldwide. Although it is remotely accessible for students using university credentials, it has serious limitations that caused a significant slowdown in the data collection process. First, the financial data available are either on quarterly bases for a maximum of 10 years period span or on a yearly aggregated level if the period exceeds 10 years. Since this study aims to analyze approximately 20 years of weekly data for each entity, it required the use of a different database, which is detailed in the next section. For this reason, BankFocus

was only involved in the selection of entities for the study. However, the second challenge presented by the database was the vast number of filtering options available and the lack of accessible online documentation to navigate them. In the end, the final criteria to filter and select the suitable banks for the dataset were the following:

1. The banks were divided into 3 groups according to the most recent 'Status' option flagged in the database:

   - ***Active companies - Active***

     The banks in this group are those marked as Active companies in the 'Active' subgroup BankFocus. They will be called the 'Active Group' or 'Active' throughout the study.

   - ***Active companies - Rest***

     The banks in this group are those marked as Active companies and belong to one of the following subgroups in BankFocus:

     – Active (administratively non-compliant)
     – Active (rescue plan)
     – Active (default payment)
     – Active (insolvency proceedings)
     – Active (reorganization)
     – Active (dormant)

     This group of banks will be called the 'Rest Group' or 'Rest' throughout the study.

   - ***Inactive companies***

     The banks in this group are marked as Inactive companies at some point in time after 2005 and belong to one of the following subgroups in BankFocus:

- Administratively suspended
- In liquidation
- Bankruptcy
- Dissolved (demerger)
- Dissolved (liquidation)
- Dissolved (bankruptcy)
- Dissolved
- Inactive (no precision)

This group of banks will be called the 'Inactive Group' or 'Inactive' throughout the study.

Each group was downloaded separately, while still maintaining all the other characteristics that follow equally.

2. The financial data concerning the selected banks, as already mentioned, were taken from a different database (Bloomberg). However, to ensure that all the entities considered were either Global ultimate owners (GUOs) or simply not minor branches or subsidiaries, which would cause to have either multiple entities with same name or lack of meaningful data, two consolidation code in BankFocus had been selected as filter:

   - Consolidated accounts with no unconsolidated companion (C1)
   - Consolidated accounts with an unconsolidated companion (C2)

   Flagging both conditions ensures to include "mother banks" that have BankFocus financial statement in one of the two categories and highly likely to have the same financial statement available in other databases, like Bloomberg, which will be detailed later.

3. In terms of geographical location, for the 'World region/Country/Region in country' flag, the following were selected:

   - United States
   - Euro Area
   - Others (not included in Euro Area): Bulgaria, Czech Republic, Denmark, Finland, Hungary, Iceland, Poland, Romania, Sweden
   - United Kingdom

   Later also some relevant Swiss entities were manually added.

4. An important additional selection to ensure the presence of relevant financial data in all the databases is:

   - Publicly listed companies
   - Formerly publicly listed companies (mostly relevant for the Rest and Inactive groups)

Once the entities were filtered, additional information for each entity was extracted such as 'Country ISO Code', 'Sifi (Yes/No)', 'BvD ID number' (unique entity number useful to connect entities data), and 'Specialisation'.

3.1.2 Bloomberg

As previously mentioned, the financial data was then extracted from the Bloomberg database using the entities' list from the above step. Bloomberg is a financial markets database that provides historical and real-time financial market and economic data, with international coverage. It also includes third-party licensed data that are not available in other databases. This database is widely used in the financial services industry. While rich in his content, the Bloomberg terminal is also expensive and most universities,

research laboratories, small firms, etc. might not have multiple accesses availability and have limits downloads per day, or simply do not consider the purchase of the software cost effective. This is one of the important practical factors to keep in mind when addressing the limits of studies on systemic risk because, under most circumstances, obtaining a considerate and reliable number of data is often very challenging. Fortunately, the necessary data for this study were easily obtained thanks to the help of the university librarian. Using as starting date the 26th of December 2003 and as ending date the 29th of December 2023 the data for each bank were collected by passing to the Bloomberg software the BvD ID numbers extracted from BankFocus. The downloaded data are divided into the following major categories:

- *Stock Metrics*
  - Market Capitalization
  - Stock Price
  - Beta

  Available on a weekly basis.

- *Financials*
  - Total Assets
  - Total Equity
  - Leverage Risk Ratio (risk-weighted assets to equity)
  - Maturity Mismatch (total loans to total deposits * 100)
  - Tier 1 ratio (tier 1 capital to risk-weighted assets)
  - Net Change of Interbank Liabilities

  Available on a quarterly basis.

- *Macro & Market Indicators (US)* Since all the subsequent calculations on the data are computationally expensive and, in some cases, complicated, the United States market was chosen as the base market. While this is likely an oversimplification, the rationale behind it lies on the assumption that the US market, being the most sophisticated and articulated one among the selected areas, drives and influences all the other markets taken into consideration, and therefore could be used as a point of reference. The US macro and market indicators are:

  - **Yield on a 10-year US Treasury bond** (USGG10YR Index)
    A rising yield can signal economic growth and higher inflation expectations, while a falling yield might suggest a weakening economy or a flight to safety by investors.
  - **Yield on a 3-month US Treasury bill** (USGG3M Index)
    This yield is highly sensitive to the Federal Reserve's policy. A higher yield often indicates that the Fed is tightening monetary policy to combat inflation, while a lower yield suggests that they are loosening it to stimulate the economy.
  - **Average yield of US high-yield corporate bonds** (BAMLH0A0HYM2EY)
    A rising yield suggests that investors are demanding a higher risk premium, indicating growing concerns about corporate creditworthiness and a less favorable economic outlook. A falling yield indicates increased confidence and more risky sentiment.
  - **S&P 500 Total Return Index** (SPXT Index)
    This index the performance of the 500 largest US publicly traded companies, including dividends.
  - **Chicago Board Options Exchange Volatility Index** (VIX

Index)

Known as the "fear index," this measures the expected short-term volatility of the S&P 500. A high VIX value indicates increased market uncertainty and investor fear. A low VIX suggests a period of market calm and stability.

Available on a weekly basis.

- **US Gross Domestic Product**

  The total monetary value of all goods and services produced within the United States

Available on quarterly basis.

All the data are structured to have the dates from 26/12/2003 to 29/12/2023 on the rows and each entity in a different column. Each single variable is in a separate Excel file that will later be converted to a CSV file. The excel files are also broken down into 3 status groups with the underscore to distinguish them and facilitate further analysis (for example, the total assets is ta_active, ta_rest, ta_inactive).

3.2 Market Variables Calculations

Market variables are used to model the time-varying nature of risk. These variables are not directly related to a bank's internal health, but rather about the broader financial and macroeconomic environment that influences how banks' assets and risks change over time. The list of the most promising choices for these market (or state) variables appears to vary among all the papers taken into account in this study. While in the theoretical sense some variables would be interesting to add, the difficulty of putting together all

the data for the calculations would be very challenging and extremely time consuming given the lack of proper resources to compute them. Some excluded variables worth mentioning are those more related to regulatory or legal issues, proposed by Weiss et al. (2014) [13] such as the Antidirector Rights Index (ADRI), which aggregates information on shareholder rights, or the Supervisory Agency Index (IoSA), which measures the quality and autonomy of the supervisory framework. Weiss et al. (2014)[13] together with De Moraes Souza et al. (2023)[14] also suggest the Herfindahl–Hirschman index (HHI) since it is a measure of high market concentration, calculated as the sum of the squared market shares of domestic and foreign banks in a country. The actual variables used in this study are listed below. All variables, but the historical risk premia and the GDP growth, have been proposed by Adrian and Brunnermeier(2011) [2]. The remaining ones are taken from De Moraes Souza et al.'s (2023)[14] work. To ensure uniformity of the unit measure, some calculations are multiplied by 100. In the end, all variables would be percentage numbers, which means that a value of 1% corresponds to the number 1.

- **Change in Credit Spread**

  Computed as the yield difference between a risky bond and a benchmark government bond:

  $$\textit{Weekly change in spread}_t = \textit{Risky bond}_t \text{ - } \textit{10Y Gov bond}_t \rightarrow$$
  $$\textit{Change in CS} = \textit{Weekly change in spread}_t \text{ - } \textit{Weekly change in spread}_{t-1}$$

  A widening spread is a signal of increasing systemic risk. It shows that investors see more risk in the corporate sector and want higher returns to compensate for potential defaults. This can lead to higher

credit risk for banks that hold corporate loans or bonds, which can result in potential losses and increased stress in financial markets, which can raise banks' funding costs.

- **Change in Yield Curve Slope**

  It measures how the difference between long-term and short-term interest rates changes over time. In this case is calculated as the difference between the 10-year and 3-month government bond yields compared to the previous period:

  $$\textit{Weekly slope}_t = \textit{10Y Gov bond}_t \text{ - } \textit{3M Gov bond}_t \rightarrow$$
  $$\textit{Change in YCS} = \textit{Weekly slope}_t \text{ - } \textit{Weekly slope}_{t-1}$$

  A flattening or inverted yield curve (when the spread between long-term and short-term rates shrinks or becomes negative) has historically been a strong predictor of economic recessions in the US. The market's demand for long-term government bonds, often seen as a safe, increases as investors become worried about future economic prospects. This drives up the price of those bonds and, consequently, pushes their yields down below short-term rates. These recessions can lead to more loan defaults for banks, financial stress, and tighter credit conditions, which can be sources of systemic risk. Conversely, a steepening curve often suggests economic recovery and lower risk.

- **Historical Equity Risk Premium**

  It measures the extra return that stocks have provided over a risk-free investment, like a government bond, over a specific period:

$$\textit{Weekly equity return}_t = \frac{S\&P500_t - S\&P500_{t-1}}{S\&P500_{t-1}}$$

$\textit{Risk free rate}_t\textit{: 10Y Gov bond}_t$

HERP = Weekly equity return$_t$ - Risk free rate$_t$

Rising risk premiums often signals that investors see a higher level of risk (like default or illiquidity) or become more risk-averse, which leads them to demand greater compensation. This situation can result in tighter financial conditions, higher borrowing costs for businesses and banks, and an increased likelihood of market instability or an economic downturn, and consequently, higher systemic risk. Conversely, a decrease in risk premiums generally indicates greater investor confidence and a lower level of systemic risk.

- **VIX**

  It is often called the "fear index" or "fear gauge" because it reflects investors' sentiment and perceived risk. A high VIX value (generally above 30) signals higher volatility, uncertainty, and investor fear. A low VIX value (generally below 20) indicates greater market stability and less overall financial stress. This variable is left as it is as no further processing is required.

- **GDP Growth**

  Both Weiss et al. (2014) [13] and De Moraes Souza et al. (2023) [14] suggest considering the country GDP growth as one of the variables as low or negative values cause higher unemployment, business failures, and widespread loan defaults (NPLs) therefore increasing systemic risk. Stable growth generally reduces systemic risk. After several tests this variable was removed as its inclusion from the models seems

to significantly improve their performances.

For all the variables calculated except for VIX, because of the lagged values, the first row will be empty, and therefore is removed from all the variables for consistency. The resulting dataset is made of 4 columns representing the state variables (excluding the GDP), and 1044 rows with the weekly dates from 02/01/2004 to 29/12/2023 as indices.

3.3 Panel Regression Variables

Adrian and Brunnermeier's (2011) [2] forward-ΔCoVaR aims to estimate the future ΔCoVaR using a panel regression of lagged variables that are considered explicative in systemic risk terms. Although more details on this methodology are detailed in the next chapter, this section illustrates the variables chosen for the regression and what drove this variables' choice in the first place.

- ***Log of Assets***

  The first chosen variable is the logarithm of assets as a proxy for size. Weiss et al. (2013) [13] and Wu et al. (2021) [11], among others, mention that taking a logarithmic transformation of asset values helps to achieve a more accurate analysis of relationships between variables by stabilizing variance, normalizing a skewed distribution, and allowing models like regression to isolate the true relationships between other variables without being overshadowed by size effects.

- ***Log of Equity***

  De Moraes Souza et al. (2023) [14] and Adrian and Brunnermeier (2011) [2] suggest using equity as a predictor of systemic risk. Following the same rationale of assets, the logarithm of equity is computed.

- ***Beta***

  Adrian and Brunnermeier (2011) [2] and Kumar et al. (2022) [10] use the entities' equity market beta to account for sensitivity. According to the authors, since a high beta means their returns value is highly sensitive to overall market movements, when a significant number of these institutions are simultaneously vulnerable, a market downturn can trigger a systemic crisis. Unfortunately, due to some issues in the data received from Bloomberg, which showed values extremely above norm, the beta was excluded from the analysis.

- ***Stock Price Returns Volatility***

  Kumar et al. (2022) [10] and De Moraes Souza et al. (2023) [14] explain that the that highly sensitive and volatile stocks can trigger panic among depositors. This panic then leads to mass withdrawals, which in turn can cause a short-term liquidity crisis, even if the bank is fundamentally sound in the long run. The liquidity crisis then can lead to a bank run to spread to other banks through contagion. Therefore, the authors believe that stock volatility is a crucial factor to monitor for regulators and banks to prevent systemic disruptions.

- ***Risk Leverage***

  This variable is taken as it is from Bloomberg. Kumar et al. (2022) [10] consider a high risk weighted assets (RWAs) over equity a good proxy of risk leverage. According to Kumar et al. risk leverage is a determinant of systemic risk, as it means that banks have a smaller capital cushion to absorb unexpected losses from their riskiest assets.

- ***Market to Book ratio (MTB) and Tobin's Q***

  These two ratios both incorporate real-time market sentiment and in-

vestor expectations about a firm's prospects. The Market-to-Book ratio is computed as standard market capitalization over book equity. López-Espinosa et al. (2012) [15] and many other scholars believe that in a crisis, a sharp drop in investors' confidence express by the MTB can signal a vulnerability to shocks, this sentiment could signal systemic risk. The standard MTB ratio however focuses solely on the equity side of the balance sheet. A more comprehensive measure of systemic risk that considers also the market's perception of a firm's overall financial health would be also to include the Tobin's Q. The Tobin's Q is a more holistic metric, calculated as the sum of the market capitalization and liabilities over assets.

- ***Maturity Mismatch***

  The closest proxy found in Bloomberg to the maturity mismatch used by Wu et al. (2021) [11] and Bellavite Pellegrini et al. (2022) [16] is: *(Total Loans / Total Deposits) * 100*. This is a good proxy for maturity mismatch because it captures the core function of banking: borrowing short-term (deposits) and lending long-term (loans). Bellavite Pellegrini et al. (2022) [16] link maturity mismatch to liquidity risk and the potential for "fire sales" of distressed assets during financial turmoil. They suggest that institutions with a lower maturity mismatch are less vulnerable to liquidity risk, which in turn reduces the risk of fire sales and the spread of risk to other institutions.

- ***Tier 1 Ratio***

  According to Varotto and Zhao (2017) [17], a higher Tier 1 ratio indicates a lower systemic risk because it signifies a stronger capital buffer, which means the bank is more resilient and better equipped to absorb

losses without failure.

- ***Net Changes of Interbank Liabilities***

  This variable was found in Bloomberg. It contains the net changes of federal funds purchased, securities sold under repurchase agreement, securities loaned, and other interbank transactions disclosed in operating activity section of cash flow statement. A large, sudden increase in interbank liabilities suggests a bank is becoming more reliant on wholesale funding markets. In a crisis, these funding sources can quickly dry up, creating a "run" on the bank and forcing it to sell assets at a loss. This variable could have been an important indicator for systemic risk analysis because it shows the extent of financial interconnectedness and a bank's liquidity reliance on other financial institutions. Unfortunately, the values of net changes of interbank liabilities on Bloomberg are almost empty for most of the considered banks and it was therefore discarded.

Chapter 4

Methodology

4.1 ΔCoVaR overview

The ΔCoVaR strategy employed follows almost entirely Adrian and Brunnermeier (2011) [2] methodology, and the exceptions will be specified. The various steps and data calculations are outlined below using numerical elements for better clarity.

1. **$\mathbf{X}_t^i$: growth rate of market valued total assets**

   Instead of using traditional accounting metrics, Adrian and Brunnermeier choose to focus on the growth rates of market-valued total asset. The market-valued total asset is calculated by multiplying a company's market equity by its leverage ratio (total assets to book equity):

   $$A_t^i = ME_t^i \cdot LEV_t^i$$

   In the case of the data considered, A_t^i is therefore calculated for each institution using the weekly market capitalization and *Total Assets/Total Equity*. This approach converts the book value of assets into a market-based figure by applying the market-to-book equity ratio. According

to Adrian and Brunnermeier, focusing on total assets is crucial because systemic risk ultimately impacts the supply of credit to the real economy. Growth is then calculated using the formula:

$$\frac{A_t^i - A_{t-1}^i}{A_{t-1}^i}$$

Once growth is computed, three datasets are generated, one for each status group (active, rest, inactive). As in the case of the market variables calculations, because of the lagged values, the first row will be empty and, therefore, is removed.

2. **$\mathbf{X}_t^{system}$: financial system returns**

   To calculate the weekly return of the financial system, Adrian and Brunnermeier (2011) [2] create a single, overall market return by taking the market-valued asset returns of all the individual financial institutions and then averaging them. The average used by the authors is a weighted average where the weights w_t^i are determined by each institution's size, specifically its lagged total assets. This means that larger institutions have a greater influence on the final system-wide return than smaller ones. The use of "lagged" assets means using the size of the institution from the previous period to do the weighting, ensuring that the calculated returns accurately reflect the real size and importance of each institution within the financial system.

   From a practical perspective, because all the data is also split in 3 status groups (active, rest, inactive), initially the sum of total assets is calculated for each row within the group and then summed again together in a final vector of sums:

   $TotalAssets_t = sum_{active} + sum_{rest} + sum_{inactive}$

The final weights are calculated dividing, for each entity i, the value of its total assets at time t to the total assets sum:

$$\frac{TotalAssets_t^i}{\sum_i TotalAssets_t}$$

Given that the reference paper uses lagged weights, before proceeding to the system returns' calculations, the last row of the dataset of the weights is removed as w_t^i will only be used up to $T-1$, where T is the number of total weeks in the sample. Removing the last row of the weights also ensures that w_t^i and X_t^i have the same number of rows and columns.

For each status group, the weights w_t^i and the total market asset returns X_t^i are multiplied with respect to entities and rows (weeks):

$$sum_t^i = X_t^i \cdot w_{t-1}^i$$

All the values resulting from the previous step are then summed over all entities and then summed over all status groups. The result of these operations is equivalent to the formula below, which assumes all the entities are in a unique data file and not divided in status groups:

$$X_t^{system} = \sum_i \mathrm{X}_t^i \cdot w_{t-1}^i \forall t$$

The resulting dataset has 1044 row values, indexed by the dates, and one single column for the financial system returns.

3. **VaR**

   The Value at Risk (VaR) is a measure of risk for financial institutions that quantifies the potential maximum loss for an institution within

a specified confidence interval. For a given probability $q \in (0,1)$, the $q\%-$VaR is the value VaR_q^i for which the probability of a loss greater than or equal to this value is exactly $1-q$:

$$Pr(X^i < VaR_q^i = q)$$

The formula shows that the VaR is the specific loss value below which a given percentage of outcomes are expected to fall. For example, a 1% VaR means that in 1% of cases, the institution's loss will be greater than or equal to the VaR value, and in 99% of cases, the loss will be less than this value.

Instead of assuming a specific distribution for asset returns (like a normal distribution), the authors use quantile regressions to estimate the relationship between an institution's returns and a set of state variables that are known to influence asset returns. This method allows to directly estimate conditional quantiles without making strong assumptions about the data's distribution.

The time-varying VaR for an institution, VaR_t^i, is calculated based on a quantile regression model. The original model regresses the institution's returns, X_t^i over lagged state variables M_{t-1} :

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

$\text{VaR}_t^i(q)$ is therefore obtained using the resulting parameters of the quantile regression for the q^{th} percentile:

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

According to Adrian and Brunnermeier, in the majority of the litera-

ture on VaR, the VaR sign is conventionally switched from negative to positive. Adrian and Brunnermeier, however, keep it negative and in this study, to remain aligned with the authors, it will be kept negative.

4. **CoVaR**

   According to Adrian and Brunnermeier (2011) [2], the prefix "Co" in CoVaR stands for conditional, contagion, or comovement. CoVaR is a systemic risk measure defined as the VaR of institution j (or the financial system) conditional on some event $\mathbb{C}(X^i)$, where $\mathbb{C}(X^i)$ denotes a specific conditioning state, most commonly that institution i is in distress at its own Value-at-Risk threshold $\mathbb{C}(X^i) = \{X^i = \text{VaR}_q^i\}$.
   CoVaR is also implicitly defined by the q-quantile of the conditional probability distribution:

$$Pr\left(X^j \leq \text{CoVaR}_q^{j|\mathbb{C}(X^i)} | \mathbb{C}(X^i)\right) = q$$

   Formally, this equation states that the conditional probability of institution j's outcome, X^j, falling at or below the loss threshold $\text{CoVaR}_q^{j|\mathbb{C}(X^i)}$, given that institution i is in state $\mathbb{C}(X^i)$, is equal to the tail probability q. In other words, $\text{CoVaR}_q^{j|\mathbb{C}(X^i)}$ represents the q-quantile of the conditional distribution of X^j given the distress of institution i.
   The CoVaR measure is more extreme than the unconditional VaR because it conditions on a "bad event" which typically lowers the conditional mean, increases variance, and can increase higher moments like negative skewness and kurtosis. CoVaR answers the question: "What is the VaR of the entire financial system if a specific institution is in distress?" by capturing the potential for distress to spread, showing how financial institutions are interconnected. A high CoVaR indicates

that the system is highly vulnerable when a particular firm is struggling.

As in the VaR case, the time-varying CoVaR for an institution, CoVaR_t^i, is calculated based on a quantile regression model. The model regresses the financial system returns X_t^{system} [1], over the institution's returns X_t^i and lagged state variables M_{t-1}:

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

$\text{CoVaR}_t^i(q)$ is therefore obtained using the resulting parameters of the quantile regression for the q^{th} percentile:

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

5. **ΔCoVaR**

   ΔCoVaR is the difference between the system's CoVaR when a firm is in distress and the system's CoVaR when that same firm is in its normal, median state:

$$\Delta CoVaR_t^i(q) = CoVaR_t^i(q) - CoVaR_t^i(50\%)$$

   This difference isolates and quantifies the marginal contribution of a single institution to overall systemic risk. A large negative ΔCoVaR

[1] López-Espinosa et al. (2012) [15] also introduce an extension to the standard CoVaR model, which they refer to as the *asymmetric CoVaR*. According to the authors, negative returns from banks, which signal a contracting balance sheet, have a much greater impact on the financial system than positive returns. The spillovers of a distressed bank can be significantly underestimated if this asymmetry is ignored. The asymmetric CoVaR model extends the standard CoVaR approach by using separate coefficients for negative and nonnegative bank returns. This allows the model to capture the system's sensitivity to a bank's returns during both periods of loss (balance sheet contraction) and periods of gain (balance sheet expansion).

means that an institution's distress significantly increases the risk for the entire financial system and it is a more effective regulatory tool than VaR because it holds firms accountable for their spillover effects, providing an incentive to reduce activities that generate systemic risk. It is important to notice that ΔCoVaR is not a measure of direct causality, but it rather captures the idea that institutions can be "systemic as part of a herd", meaning if a group of small firms are all exposed to a common risk factor, the distress of one firm signals that others are likely also struggling, increasing system-wide risk.

4.2 Implementation

The model implementation was both computed using Linear Quantile Regression and Neural Network Quantile Regression. Adrian and Brunnermeier (2011) [2] also mentions that CoVaR can be estimated using a GARCH model. This is an alternative approach that models the time-varying covariance between institutions and the financial system. The GARCH method, while more complex and requiring strong distributional assumptions, produces estimates of ΔCoVaR that are quite similar to those derived from the quantile regression method, confirming the validity of the paper's primary approach.

In both linear quantile regression (LQR) and neural network quantile regression (NNWR) implemented models, $q = 0.05$ is chosen instead of the more extreme q = 0.01. According to Adrian and Brunnermeier, the results are largely qualitatively similar for both the 1% and the 5% quantiles. However, from a practical standpoint, the 5% quantile is easier to implement because the tail risk measures suffer from the rarity of "tail observations": the 5% level provides more data points for estimation, making it a more robust and

statistically reliable measure than the 1% quantile, which relies on a smaller number of extreme events.

All implementation steps are performed using a Google Notebook Colab because it provides a virtual machine on Google's cloud servers with more RAM than a personal computer, bypassing local hardware limitations and allowing the execution of intricate and computationally expensive calculations that will be discussed later in this chapter.

The implementation process started with **preprocessing** of the data, since the initial raw data from Bloomberg required some preliminary refinements. To avoid errors or inconsistencies in later stages of the code, characters such as "<null> " and " #N/A N/A " which signal lack of available data for specific entities on specific dates in the Bloomberg terminal are simply converted to "nan" (not a number) values.
Since companies of all kinds, including banks, usually do not issue weekly financial statements due to the impracticality and high cost of preparing such frequent reports and because they are also not required by regulatory bodies to do so, the financial data from Bloomberg, with the exclusion of the stock data, are provided on a quarterly basis. However, to follow Adrian and Brunnermeier's (2011) [2] strategy, the preferred level of granularity is weekly. Some interpolation on the quarterly data is therefore necessary to proceed.
Adrian and Brunnermeier (2011) [2] specify that they "generate weekly leverage data as a linear interpolation of the quarterly data. The interpolated values of the leverage are computed only between no more than two consecutive quarters of the missing leverage data" (p. 1). López-Espinosa et al. (2012)

[15] also mention interpolation and state: "we assume that the leverage ratio remains (approximately) constant throughout successive weeks within any given quarter/semester, thereby approaching the unobservable weekly value with the low-frequency data available in the period" (p. 272). He further notes that an alternative smoothing method, "a well-known technique in applied finance," is the use of "cubic spline interpolation," though he concludes that "the final results are not sensitive to this consideration" (p. 272).
Adrian and Brunnermeier state that assuming that the leverage ratio as Assets over Equity could be directly interpolated or that it remains constant throughout successive weeks within any given quarter/semester. However, to increase the precision of interpolation and avoid unrealistic jumps in the data, Total Assets and Total Equity are interpolated separately. Both balance sheet components will also be used on their own, and therefore this strategy allows consistency in future calculations. Following Adrian and Brunnermeier's method, the interpolation implemented works under two conditions:

- For short initial quarterly gaps (one or two missing quarterly values) where a valid preceding data point exists, linear interpolation is applied to fill the corresponding weekly data points. This means estimating values along a straight line between the last known point and the first available point in the series.

- If the gap is larger than two quarters or if no preceding valid data point is available, no interpolation is performed for that specific gap; only the existing valid quarterly data points are placed onto their corresponding weekly dates, with the weekly dates in between remaining unfilled.

This strategy is applied for all the financial data available on quarter basis. The resulting datasets have a total of 1045 rows with weekly dates as

indices and an initial amount of 632 columns representing banks for the Active group, 30 columns for the Rest group, and 622 columns for the Inactive group.

An important preliminary mention to make in the quantile regression methodology is the **pinball loss function**. The pinball loss function, also known as the quantile loss function, is a tool used in quantile regression to measure the accuracy of a predicted quantile. It is asymmetric, which means that it penalizes overestimation and underestimation differently, depending on the chosen quantile, q. The function is defined as follows:

$$L_q(y, \hat{y}) = \begin{cases} (y - \hat{y}) \cdot q & \text{if } y \geq \hat{y} \\ (\hat{y} - y) \cdot (1 - q) & \text{if } \hat{y} > y \end{cases}$$

Here, y is the actual value and $\hat{y}$ is the predicted quantile. Quantile regression aims to minimize this loss. For a low quantile such as $q = 0.05$, the penalty for underestimation is very small (weighted by 0.05), while the penalty for overestimation is much larger (weighted by 0.95). For a 5% VaR or CoVaR model, the goal is to try to capture the worst event out of 100. The model will underestimate the loss for approximately 95% of the data points and will overestimate the loss for approximately 5% of the data points. Although the penalty per observation for underestimation is small, the number of times it happens results in a large cumulative penalty. Conversely, the penalty for overestimation is large, but it happens very rarely, therefore the model must balance these two.

Moving on to the actual implementation in Python, Adrian and Brunnermeier (2011) [2] cite incremented **expanding windows**. This approach, also known as *expanding window analysis*, progressively increases the size of the

sample set by adding new data points while keeping the starting point fixed. To keep both the linear model and the neural network model consistent and because using large expanding windows instead of small ones is generally better for training neural networks on time series data since it allows the model to learn from a larger more complete history, the expanding windows will be incremented by 260 weeks, approximately 5 years. This is what Keilbar and Wang (2019) [9] suggest as the minimum number of observations for the model to achieve convergence.

The implemented expanding window algorithm works to iterate on all entities and checks if the first 260 observations for the first window are not NA. If they are, that entity will be temporarily skipped; if not, the quantile regressions and ΔCoVaR will be performed. The next window will include the previous window plus another 260 observations, for a total of 520 observations considered. For each entity, if at least 260 observations over the 520 are not NA, the quantile regression is performed, filling also the NA values from the previous calculations if the entity was skipped, but not overwriting the previous calculations if the values for that entity were filled in a previous window. The algorithm proceeds this way for all remaining windows.

4.2.1 ΔCoVaR Models

The linear model computes all the quantile regression detailed in the ΔCoVaR overview section over expanding windows as previously explained. The maximum iteration parameter is set to 10000 to ensure that the model attempts convergence without overloading the processing power. The difference between this study and Adrian and Brunnermeier's (2011) [2] is that the same steps are also computed using neural networks.

A **neural network** is a computational model inspired by the human brain.

It consists of interconnected nodes, or "neurons," organized into layers: an input layer, one or more hidden layers, and an output layer. Each connection between neurons has a numerical weight, and each neuron has a bias. When a neural network receives input data, it processes it through its layers, and each neuron in a hidden layer performs a two-step calculation. First, it computes a weighted sum of the inputs from the previous layer, adding its bias to the sum. Second, it applies an activation function to this sum. The activation function introduces non-linearity, allowing the network to learn complex patterns, and the output of this function is then passed as input to the next layer. This process continues until the data reach the output layer, which produces the prediction. In supervised learning, the network evaluates its error against the true values, calculates gradients via backpropagation, and updates its weights and biases using an optimization algorithm to minimize the loss.

For a single neuron, the calculation is the weighted sum of its input plus a bias:

$$z_j^{(l)} = \sum_i w_{ji}^{(l)} a_i^{(l-1)} + b_j^{(l)}$$

where:

- j is the the neuron in layer l
- $a_i^{(l-1)}$ is the activation (output) of neuron i in the previous layer $l-1$
- $w_{ji}^{(l)}$ is the weight of the connection from neuron i in layer $l-1$ to neuron j in layer l
- $b_j^{(l)}$ is the bias of neuron j in layer l

The final output of neuron j is then obtained by applying the activation function f to this weighted sum:

$$a_j^{(l)} = f(z_j^{(l)})$$

The backpropagation algorithm to update the weights and biases to minimize the loss function is as follows:

$$w_{new} = w_{old} - \alpha \frac{\partial L}{\partial w}$$

where α is the learning rate and $\frac{\partial L}{\partial w}$ is the gradient of the loss with respect to the weight.

According to Taylor (2000) [3], neural network quantile regression (NNQR) is a technique that uses a neural network to model the relationship between input features and a specific quantile of the target variable's distribution. The core of this method lies in training the network to minimize the pinball loss function. From a theoretical point of view, neural networks are superior to linear methods as they can capture highly complex, nonlinear relationships and dependencies in the data that are essential for accurately predicting extreme tail events like VaR.

The chosen parameters for this analysis are the following:

- Hidden layers: 2
- Nodes per layer: 64
- Activation function: ReLu $\rightarrow f(z) = \max(0, z)$
- Regularization (which prevents overfitting by adding a penalty term to the loss function):

  - L1 Regularization (Lasso) $\rightarrow$ Penalty Term $= \lambda \sum_{i,j} |w_{ji}|$ with λ: 0.001
  - L2 Regularization (Ridge) $\rightarrow$ Penalty Term $= \lambda \sum_{i,j} w_{ji}^2$ with λ: 0.001

- Dropout rate (fraction of neurons to drop during training to avoid them to co-adapt too much): 0.1
- Epochs (number of times the model will go through the entire dataset): 50
- Batch size (number of training samples processed before the model's internal parameters, weights and biases, are updated): 32
- Patience (if the model's performance on a validation set does not improve for a number of patience epochs, the training process is stopped to prevent overfitting): 10
- Validation split ratio (the fraction of the training data to be set aside for validation): 0.2

The choice of the number of hidden layers and nodes per layer makes it a moderate complex network, without too much risk of overfitting while retaining nonlinear relationship capturing capacity. The ReLu activation is kept standard, just as the Lasso and Ridge regularization. Training for 50 epochs means that the network will make 50 full passes through the training data, and a batch size of 32 means that the model will update its parameters after processing 32 samples at a time. The patience parameter set to 10 tells the model to stop training if its performance in the validation data does not improve for 10 consecutive epochs. This prevents the model from continuing to train and overfit to the training data after it has already found an optimal

solution for the validation set.

Implementing a neural network for each entity as this analysis requires, presents significant computational challenge: the process is extremely expensive, exceeding the capabilities of a standard local machine. For a sample of thousands of total institutions and thousands of weeks of data, running a network for each entity on Google Colab notebooks required the purchase of more than 300 computing units and hours of running time, making the process infeasible without a high-performance computing environment.

To perform some of the performance evaluations on the two quantile regression models, the **pseudo-squared R** will be used. This measure is used to evaluate the explanatory power of a statistical model, and it is primarily used to compare different versions of the same type of model, in this case quantile regression, on the same data to see which one performs better. The pseudo R-squared is equivalent to the standard R-squared used in ordinary least squares (OLS) regression, which indicates how much of the unexplained variation in the dependent variable is accounted for by the inclusion of predictors in the full model, relative to the baseline model. Because models like logistic or quantile regression do not optimize for least squares and their residuals do not meet the same assumptions, the standard R-squared could not be used.

The formula for calculating the pseudo R-squared is:

$$R^2_{pseudo} = 1 - \frac{L_{full}}{L_{intercept}}$$

where:

- L_{full} represents the error, in this case the mean pinball loss, of the full

and more complex model, either linear quantile regression (LQR) or neural network quantile regression (NNQR), measuring how well the model's predictions align with the actual data.

- $L_{intercept}$ is the error, in this case the mean pinball loss, of the "intercept-only" or "null" model, which is the baseline model. This model makes predictions based only on a single value, in this case the specified quantile of the target variable), without considering any predictor variables.

The formula therefore compares the loss of the full model with the loss of this simple baseline mode. If the complete model is a good fit, its loss will be much smaller than the loss of the baseline model, causing the ratio to be close to zero and the pseudo-squared value to be close to 1.

An important observation must be made in the case of CoVaR calculations, where the dependent variable X_t^{system} is structured to have a single identical value for all entities at any given time. This characteristic might lead to having certain entities with a high ability to explain the data, while others, for reasons that might be related to the quality of the data or the presence of more 'noisy' observations, will not achieve a good pseudo-R-squared. When looking at the pseudo R-squared frequencies plots, the limitations of using a single dependent variable for all entities is an important aspect to consider.

4.2.2 Forward-ΔCoVaR

Adrian and Brunnermeier (2011) [2] developed a method to address the procyclicality issues that contemporaneous risk measures face. In simple words, the authors found a way to address the fact that risk tends to build up before the actual financial crisis occurs, and contemporaneous risk measure estimates would be unnecessarily tight after adverse events and unnecessarily loose in periods of stability. Their forward-ΔCoVaR measure overcomes

this "procyclicality pitfall" by using a firm's current, observable balance sheet characteristics to forecast its systemic spillover risk (ΔCoVaR_i) several quarters ahead, ensuring that risk warnings are triggered during credit booms rather than after market distress has already materialized. Methodologically, Adrian and Brunnermeier implemented the forward-ΔCoVaR by running a panel regression of time-varying ΔCoVaR on a vector of lagged institution-specific characteristics. This framework directly maps a firm's current structural profile to its expected future contribution to systemic distress. The variables selected for this study are mentioned in the previous section.

Before effectively computing the panel regression, a preliminary correlation matrix was made on the independent variables to verify any possible correlation that could have misled the calculations. After the analysis on the correlation matrix, the first step of the process is to convert the weekly ΔCoVaR calculated both with LQR and NNQR into quarter values by adding 12 or 13, depending on the quarter and the year, weekly values. If at least one of the values is "nan" (not a number), the whole quarter is set as "nan". The independent variables for the regression are cut of the last quarter, for the quarter lag, the last year, for the single year lag, and the last two years, for the two-year lag. The opposite operation is performed for all ΔCoVaR values by removing the first quarter, the first year, and the first two years, respectively. In doing so, for each status group, there are six sets of data: two types of ΔCoVaR, one with $q = 0.05$ and one with $q = 0.5$, each with three different lag options. The next key step before the panel regression is to convert all the variables to the 'long format'. In the long format, each row represents a unique observation. This is done by stacking all the entities one below the other so that the first column of the new data set refers to

the quarter, the second refers to the name of the entity, and the third is the name of the variable considered. In the end, there will be as many datasets in long format as the number of variables taken into consideration. Once this step is completed, all values are merged by entities and by date, and finally the panel regression is run.

When computing a panel regression, it is possible to control for entity fixed effects and time fixed effects, a standard method in econometrics for robust analysis of panel data. Entity fixed effects account for any unobserved characteristics of each entity that might be correlated with the independent variable and do not change over time like, for example, control for management style or corporate culture, which are time-invariant but could affect its performance. The panel regression controls for entity fixed effects by essentially creating a unique intercept for each entity.

Time fixed effects, on the other hand, account for fixed effects for each period as it controls for any unobserved, entity-invariant shocks. These shocks affect all entities each year or period, such as macroeconomic recession or a new regulatory policy that would affect all banks in a specific year. The time fixed effect parameter helps control for that common shock. When both entity fixed-effects and time fixed-effects are set to True, the Python function estimates a two-way fixed-effects model, which is mathematically represented as:

$$y_{it} = \alpha_i + \gamma_t + \beta^{'} x_{it} + \epsilon_{it}$$

where:

- y_{it} is the dependent variable for entity i at time t
- α_i is the unobserved time-invariant individual effect (entity effect)

- γ_t is the unobserved entity-invariant time effect
- β is the vector of the coefficient of the independent variable
- x_{it} is the vector of the independent variables for entity i at time t
- ϵ_{it} is the error term

This function is particularly useful for reducing omitted variable bias when there is a chance that unobserved characteristics are correlated with the variables of interest. The final coefficients are derived by using a technique called the "within transformation", also known as *demeaning*, which eliminates the unobserved fixed effects (α_i and γ_t) from the equation, allowing for a consistent estimation of the coefficients (β).

The demeaning process subtract to each variable its own average at each point in time:

$$y_{it} - \bar{y}_i = (\alpha_i - \bar{\alpha}_i) + (\gamma_t - \bar{\gamma}_i) = +\beta'(x_{it} - \bar{x}_i) + (\epsilon_{it} - \bar{\epsilon}_i)$$

$$y_{it} - \bar{y}_t = (\alpha_i - \bar{\alpha}_t) + (\gamma_t - \bar{\gamma}_t) = +\beta'(x_{it} - \bar{x}_t) + (\epsilon_{it} - \bar{\epsilon}_t)$$

Since the fixed effects are constant, when subtracting from the fixed effects their own average, the fixed effects coefficients are eliminated from the equation: $\bar{\alpha}_i - \alpha_i = 0$ and $\bar{\gamma}_i - \gamma_i = 0$.

The PanelOLS function then combines the new transformations into a single step which results in a double-demeaned variable:

$$y_{it}^{**} = \beta^{'} x_{it}^{**} + \epsilon_{it}^{**}$$

where the double-asterisks ($**$) denote the double-demeaned variables. The final coefficients for beta are then estimated by running a standard Ordinary Least Squares (OLS) regression on these transformed (demeaned) variables.

Chapter 5

Performance Comparison

After computing the VaR and CoVaR for both $q = 5\%$ and $q = 50\%$, and with both LQR and NNQR, a quantitative comparison of the models' performance is shown below. For spatial reasons, the graphs reported in this chapter are specific to the active status group. The rest group, which is small compared to the other status groups, has been merged to the inactive group. The graphs of the merged group display very similar performance results to the active group, and therefore the following conclusions could be considered valid across all datasets. Furthermore, again for spatial reasons, the performance of CoVaR in the median state $q = 50\%$ is also not shown here, but the performance levels of the models were consistent with the behavior of CoVaR in the low quantile $q = 5\%$.

5.1 Pinball Loss Function

The histograms in 5.1, 5.2, 5.3, and 5.4 show the frequency distributions of the pinball loss function (defined in a previous chapter) for the LQR model and the NNQR model. In the case of $\text{VaR}_{0.05}$ calculations in 5.1 and 5.3, both the linear and neural networks show a highly skewed distribution of pinball

losses, with a high frequency of values concentrated near zero. This pattern indicates that these models perform with a small error most of the time, with a very low number of predictions resulting in larger losses. However, in the case of the $\text{CoVaR}_{0.05}$ calculations shown in 5.2 and 5.4, both the linear and the NN cases display a more varied distribution of pinball loss, lacking the concentration of losses near zero and showing peaks at nonzero values instead. This indicates a larger distributed range of prediction errors. The similarity in the distributional shapes between the linear and NN pinball losses, both for $\text{VaR}_{0.05}$ and $\text{CoVaR}_{0.05}$, implies that the performance of the two types of models is very similar in terms of pinball loss.

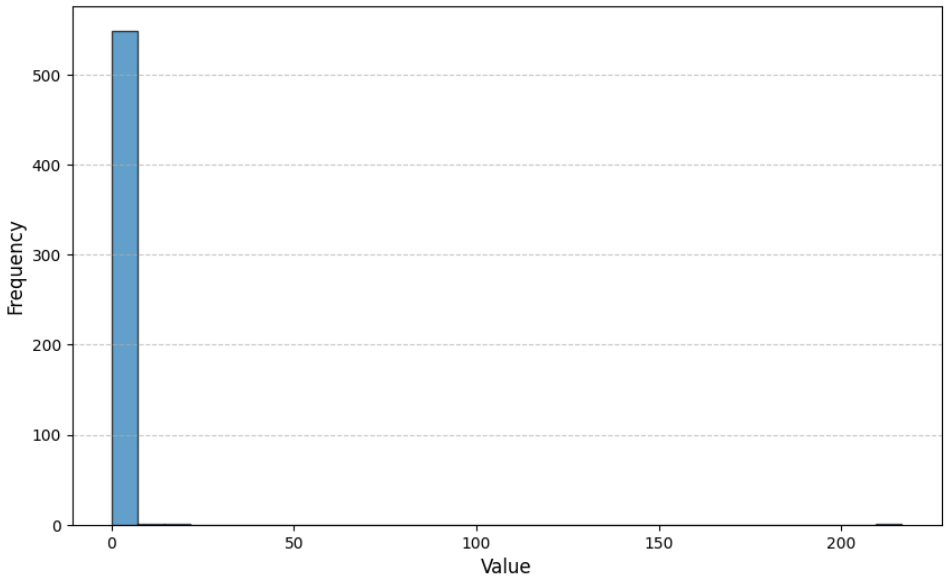


Figure 5.1: Pinball loss frequency histogram of 5%VaR with LQR (Active)

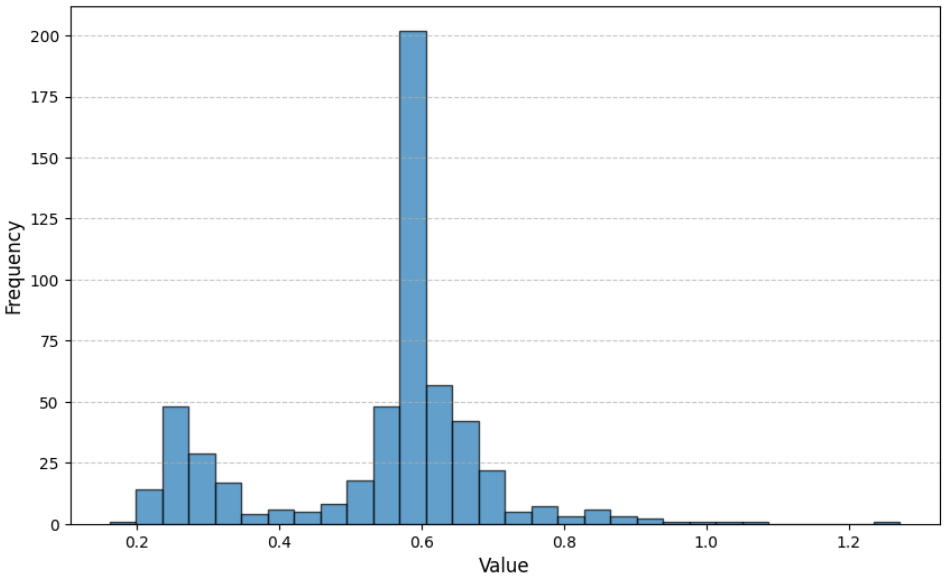


Figure 5.2: Pinball loss frequency histogram of 5%CoVaR with LQR (Active)

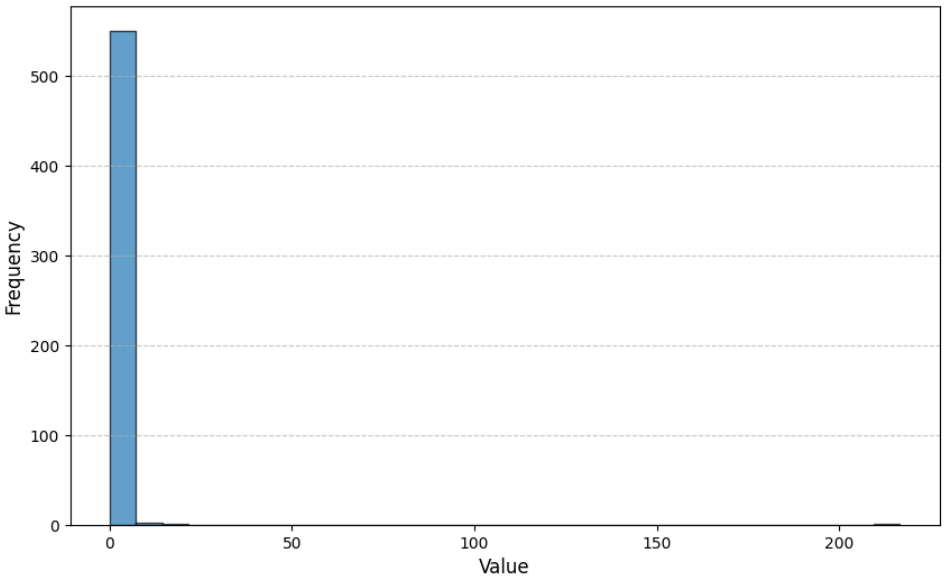


Figure 5.3: Pinball loss frequency histogram of 5%VaR with NNQR (Active)

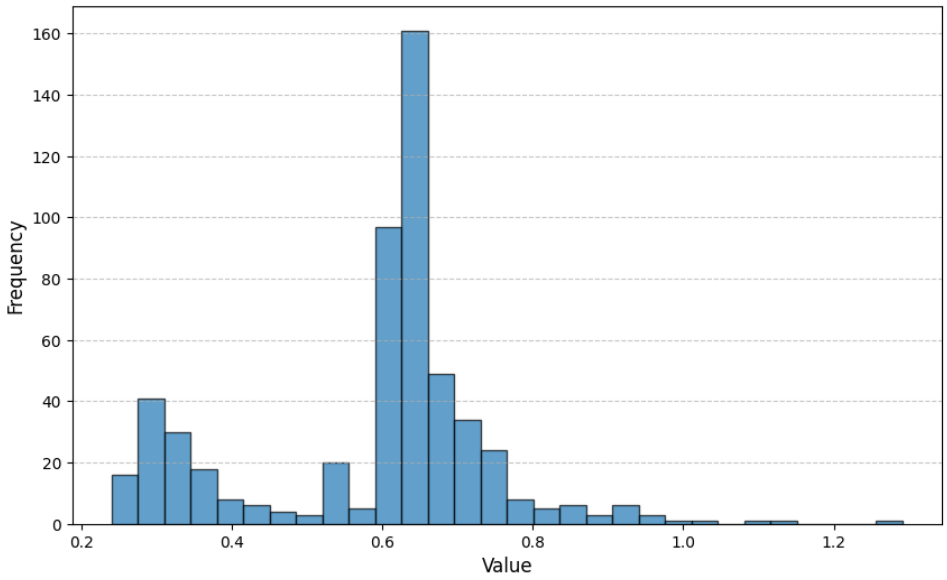


Figure 5.4: Pinball loss frequency histogram of 5%CoVaR with NNQR (Active)

5.2 Negative Residuals

For a quantile regression model at a low quantile, such as $q = 0.05$, the goal is to produce predictions where approximately 5% of the actual observations fall below the predicted value. As shown in the methodology, a negative residual $y_i - \hat{y}_i < 0$ occurs when the predicted value $\hat{y}_i$ is higher than the actual value y_i. Therefore, the ideal distribution of negative residuals would show a concentration around this target percentage because a negative residual indicates an over-prediction, and for a low quantile like the 5th percentile, a limited number of this type of errors is expected. As already mentioned in the methodology, the pinball loss function is designed to enforce this by highly penalizing over-predictions (negative residuals) compared to under-predictions. A model that consistently produces a low percentage of negative residuals is therefore considered to be better aligned with the objective of

the quantile regression.

Based on the histograms of negative residuals for the $\text{VaR}_{0.05}$ and $\text{CoVaR}_{0.05}$ calculations in 5.5, 5.6, 5.7, and 5.8 the NNQR models outperform the linear ones. The LQRs for both $\text{VaR}_{0.05}$ and $\text{CoVaR}_{0.05}$ in 5.5 and 5.6 consistently show a peak in their negative residual distributions around $5-6\%$, indicating that they often overestimate the actual value. On the other hand, the NNQR models in 5.7, and 5.8 for both calculations show peaks around 5% for VaR and between 2.5% and 4% for CoVaR. This lower percentage indicates that when these NN models do overpredict, they do so by a significantly smaller margin than the linear models, a fundamental advantage of NNQR over the LQR.

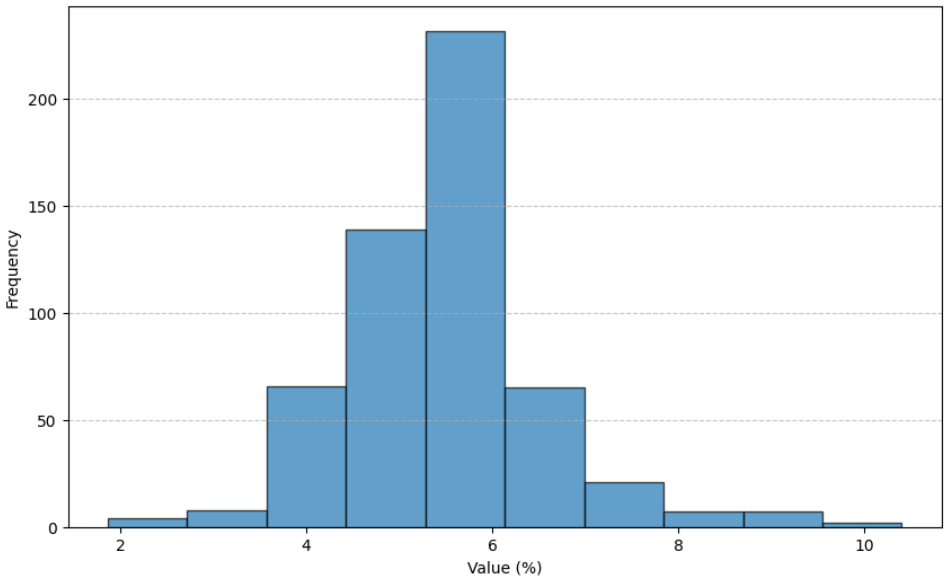


Figure 5.5: Negative residuals percentage of 5%VaR with LQR (Active)

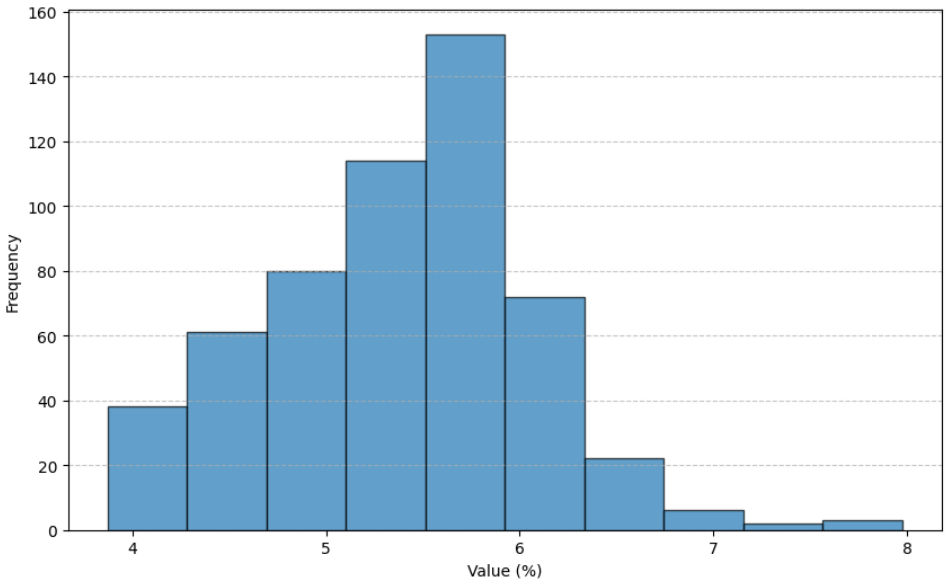


Figure 5.6: Negative residuals percentage of 5%CoVaR with LQR (Active)

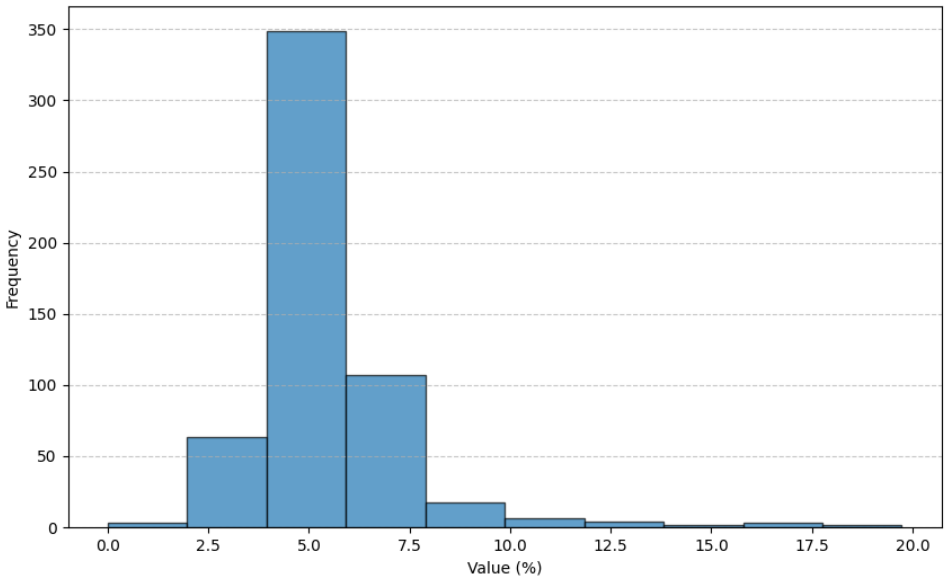


Figure 5.7: Negative residuals percentage of 5%VaR with NNQR (Active)

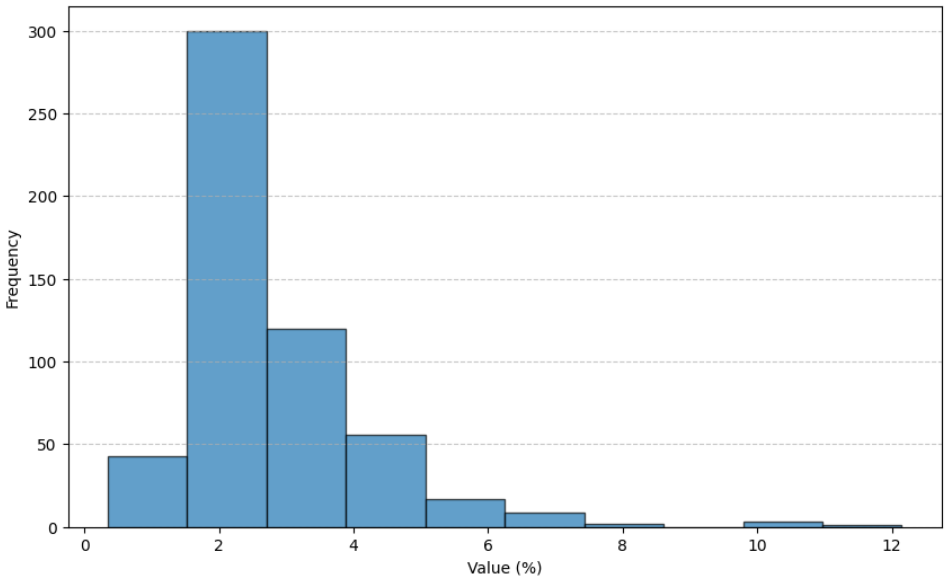


Figure 5.8: Negative residuals percentage of 5%CoVaR with NNQR (Active)

5.3 Pseudo R-Squared

As outlined in the methodology, a higher pseudo R-squared signifies that the model more accurately captures the extreme tail relationship between the individual institution's losses and system-wide distress. The distributions of pseudo-R-squared values shown in 5.9, 5.10, 5.11, and 5.12 reveal an important difference in model performance between the $\text{VaR}_{0.05}$ and $\text{CoVaR}_{0.05}$ calculations, with less variation observed between the linear and neural network models themselves. The CoVaR models with both LQR and NNQR, display in 5.10 and 5.12 a bimodal distribution of pseudo-R-squared values, with a peak around $0.1 - 0.2$ and a secondary smaller peak at $0.5 - 0.6$. This behavior indicates that the models' explanatory power for CoVaR is inconsistent, ranging from a low to a significantly better fit. On the other hand, the VaR models in 5.9 and 5.11 show a more consistent performance, with a

single peak distribution centered on 0.1 for both types of model. The values for VaR are more narrowly concentrated, suggesting a more stable, albeit modest, level of fit. The similarity between the linear and neural network models within each calculation type implies that the intrinsic characteristics of the VaR and CoVaR data are the primary determinants of the pseudo R-squared distributions, rather than the specific model.

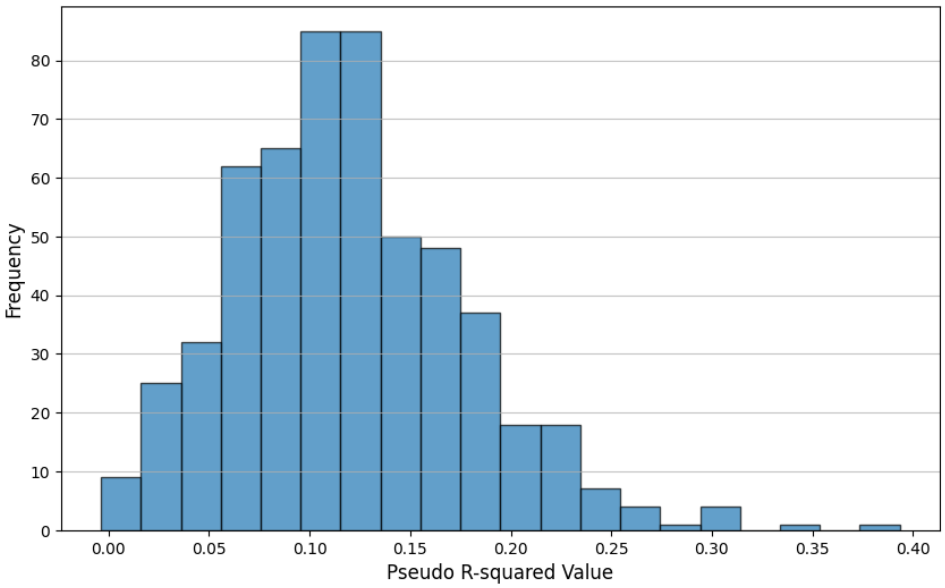


Figure 5.9: Pseudo R-squared of 5%VaR with LQR (Active)

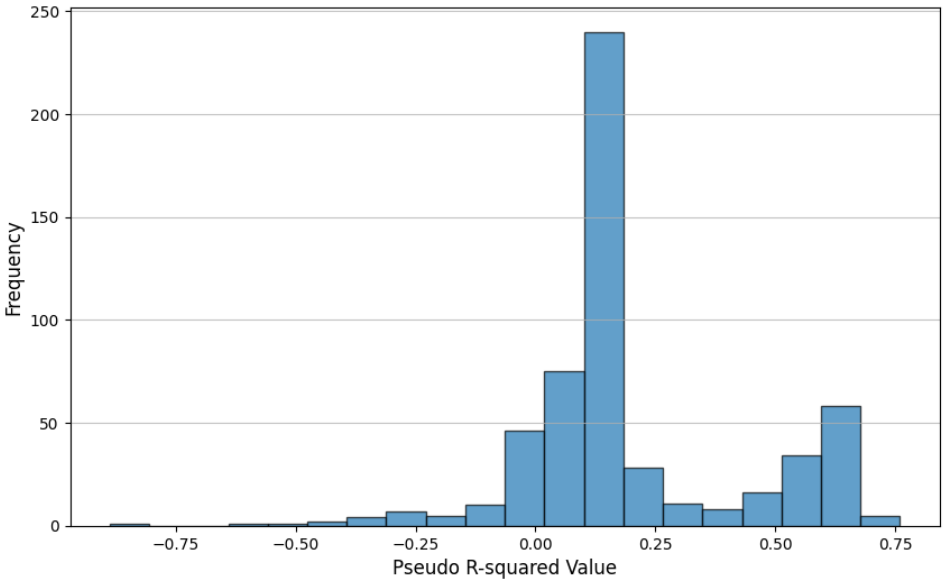


Figure 5.10: Pseudo R-squared of 5%CoVaR with LQR (Active)

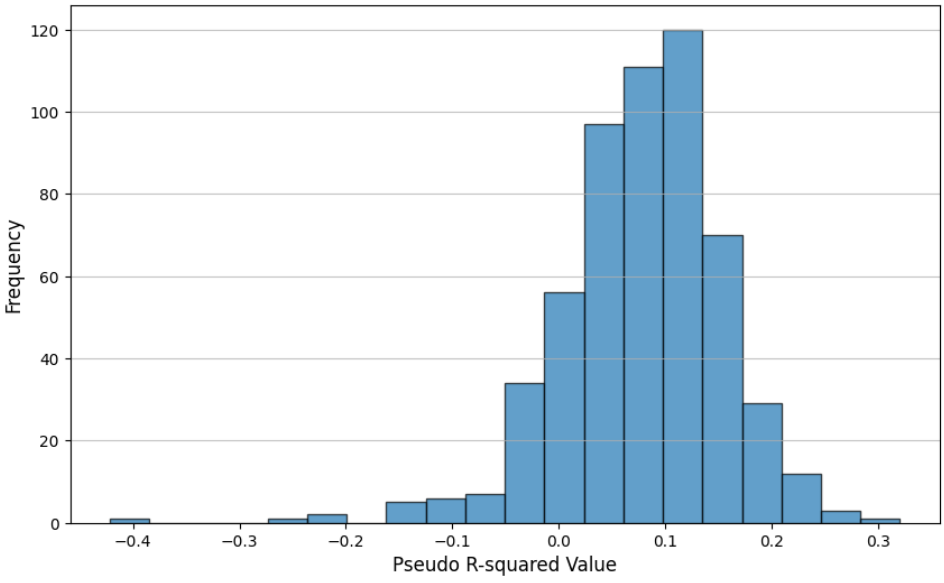


Figure 5.11: Pseudo R-squared of 5%VaR with NNQR (Active)

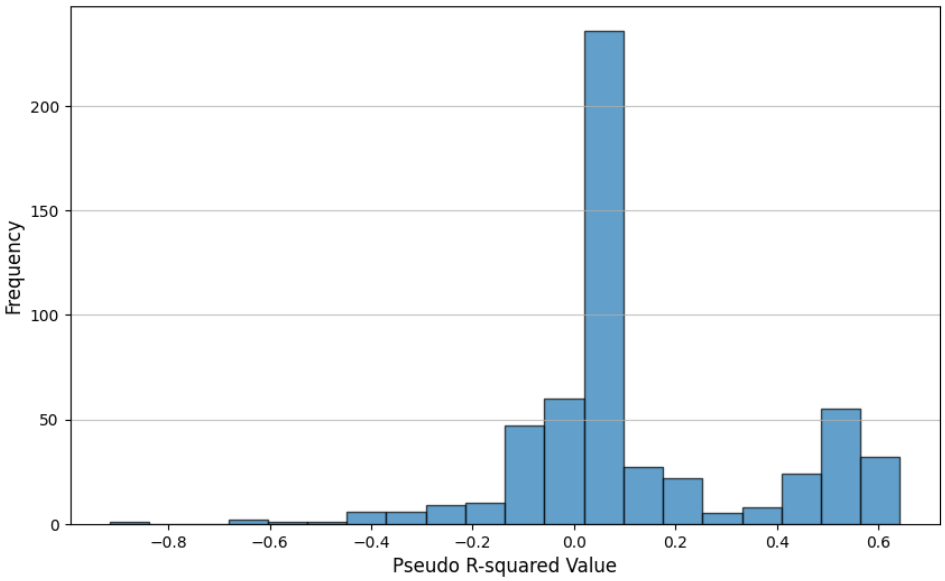


Figure 5.12: Pseudo R-squared 5%CoVaR with NNQR (Active)

5.4 Conclusion

The analysis of the performance of the linear quantile regression (LQR) and neural network quantile regression (NNQR) models across several metrics reveals that there is no single definitive best overall model. The distributions of the pinball loss function and the pseudo-R-squared values indicate a generic similarity between the two types of model, implying that the intrinsic characteristics of the input data used for VaR and CoVaR are the major contributor to the performance levels. A more distinct difference emerges in their percentage of negative residual plots, where NNQR seems to perform best.

An important observation to be made is that the CoVaR calculations present a difficult modeling challenge since the dependent variable, the financial system returns, is a single range of values for each date, kept equal for all entities.

This is likely the cause for the bimodal pseudo R-squared distributions in both models. While this might lead to considering the pseudo R-squared distributions comparison inconsistent, the NNQR still can be considered to be the slightly better performing model due to the fact that it achieves a lower percentage of negative residuals in both VaR and CoVaR, which is a good result especially for CoVaR, considering the mentioned issues.

Chapter 6

Graphical Analysis

This analysis involves plotting Var, CoVaR, and ΔCoVaR for some of the entities in the dataset. Given the magnitude of the data used in this study, each one of the entities in the three status groups (active, rest, inactive) is also subgrouped within the main status group by both bank specializations and geographical area. The purpose of the analysis is to provide a more qualitative comparison of LQR and NNQR for the calculation of Var, CoVaR, and ΔCoVaR.

6.1 Entire period

In the initial analysis, the Var, CoVaR and ΔCoVaR plots are made for the entire time series considered in this study. The figures below show some examples of graphs from commercial banks belonging to the active group both in the United States 6.1 and the Euro Area (6.2, and from US bank holdings belonging to the rest group 6.3 and the inactive group 6.4. For spatial restrictions, all other graphs are not reported. The plots of Var, CoVaR and ΔCoVaR calculated with NNQR are also not displayed for the same reason. However, the plots of the NNQR version of the risk metrics, at least in the

entire period considered, appear to be extremely similar to those from LQR. Across all plots (even those not included) for each status group, geographic area, and specialization, it emerges that there are synchronized drops in the three systemic risk measures VaR, CoVaR, and ΔCoVaR, during periods of financial stress. The most notable ones, from the entire time series of data, appear to be the financial crisis in late 2008 and early 2009 and again in late 2020, likely due to the COVID-19 pandemic. These synchronized drops indicate that during crises, financial institutions across all sectors become highly correlated, and their risk measures, particularly their contribution to systemic risk, increase significantly in absolute value. The VaR, which measures the individual risk of each institution, becomes more negative for almost all institutions, signaling a dramatic increase in individual risk. Similarly, the large drops in CoVaR highlight that an institution's distress is associated with a larger potential loss for the system, underscoring their systemic importance during these periods. ΔCoVaR , which measures an institution's marginal contribution to systemic risk, becomes significantly more negative for all sectors during the 2008 and 2020 crises. This demonstrates the "tail co-movement" that is a central feature of systemic risk.

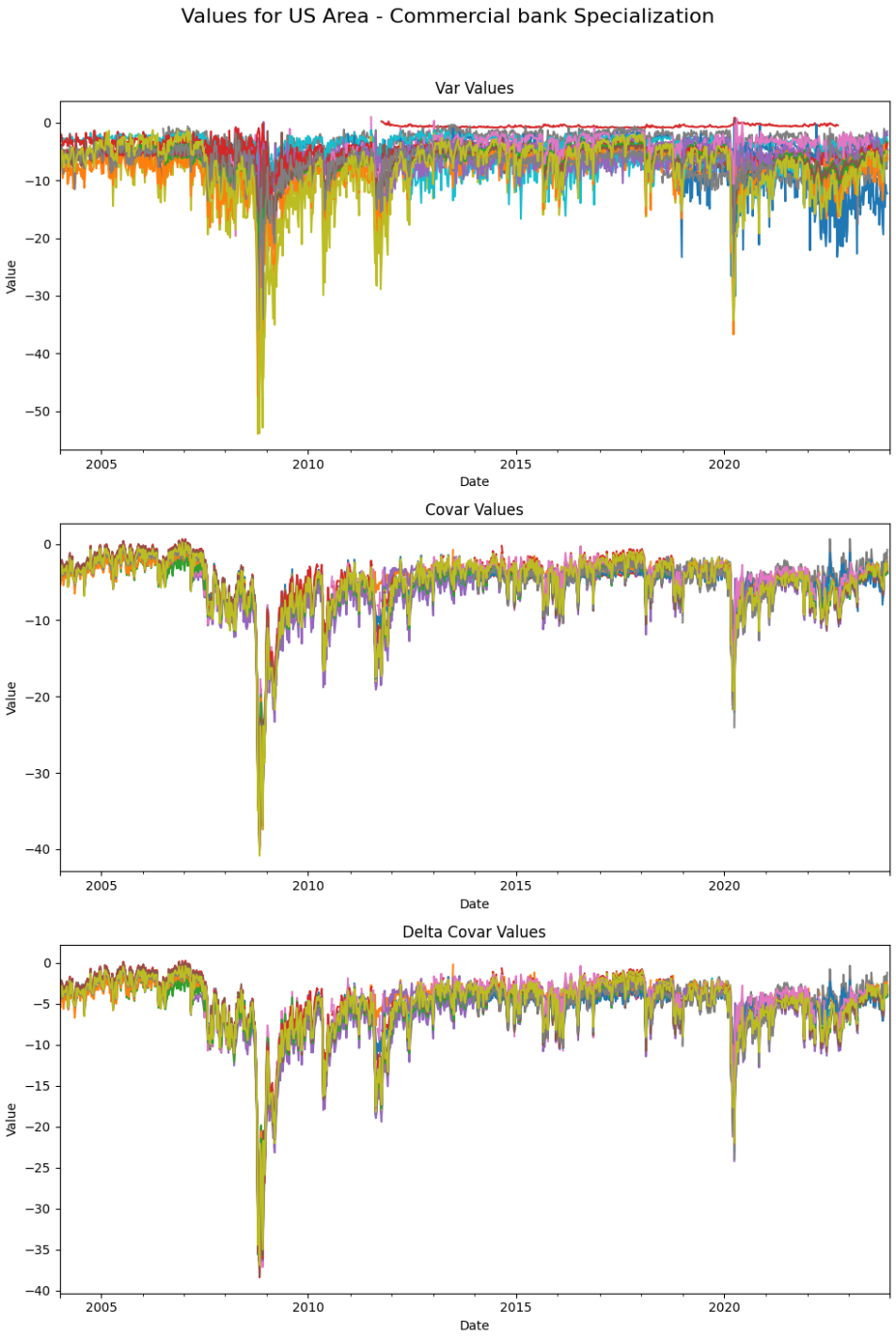


Figure 6.1: Var, CoVaR, ΔCoVaR for commercial banks - US area (Active)

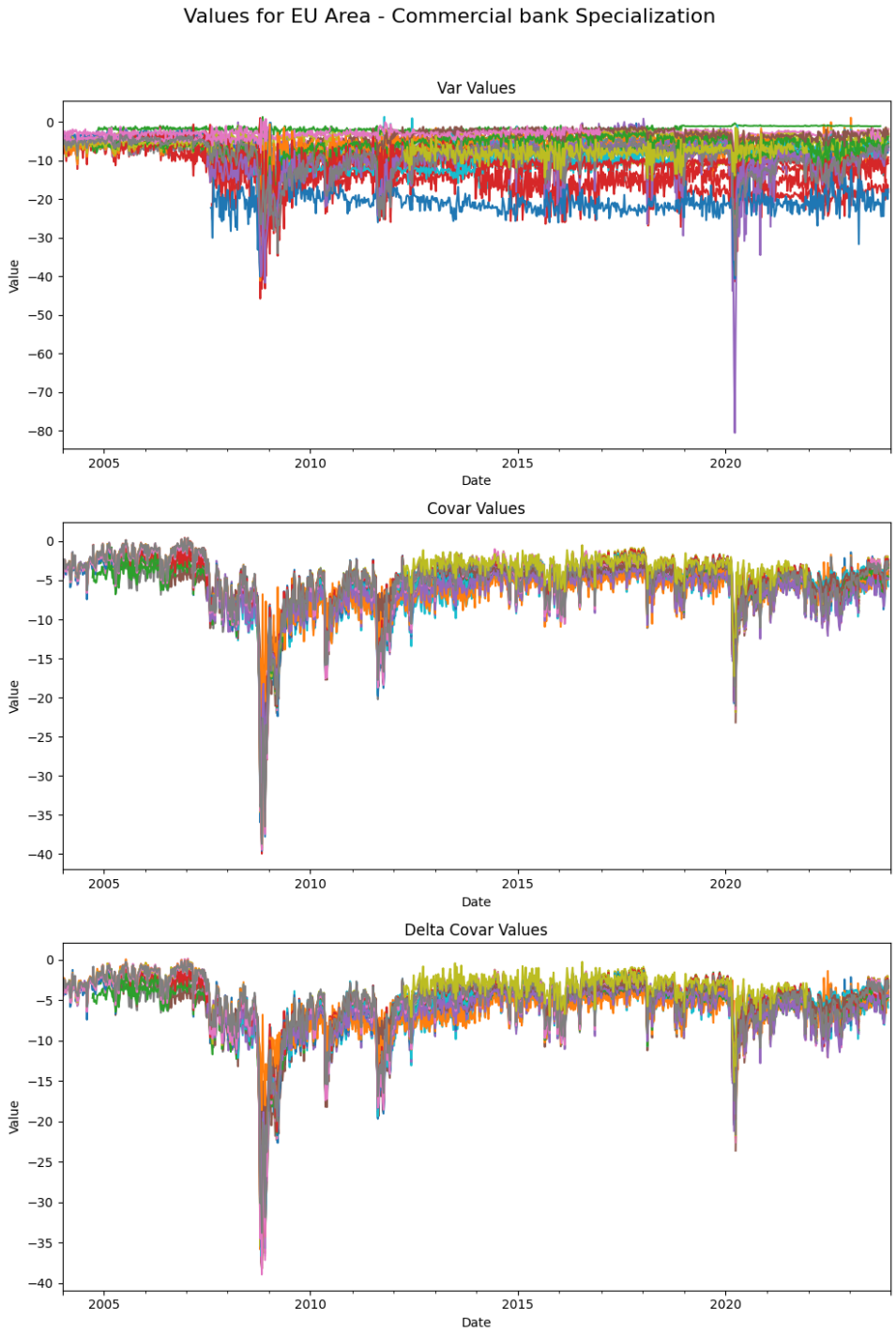


Figure 6.2: Var, CoVaR, ΔCoVaR for commercial banks - Euro area (Active)

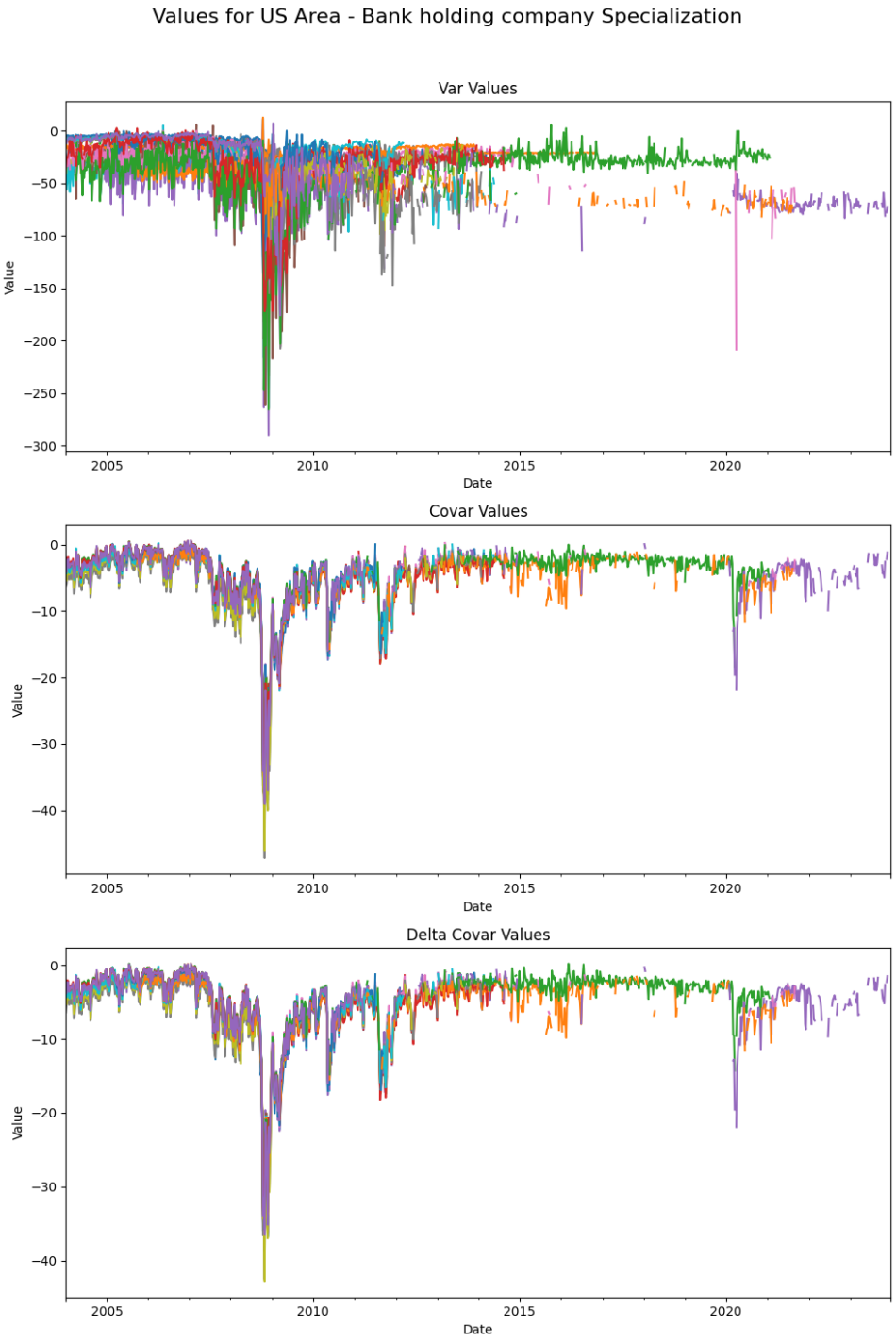


Figure 6.3: Var, CoVaR, ΔCoVaR for bank holding companies - US area (Rest)

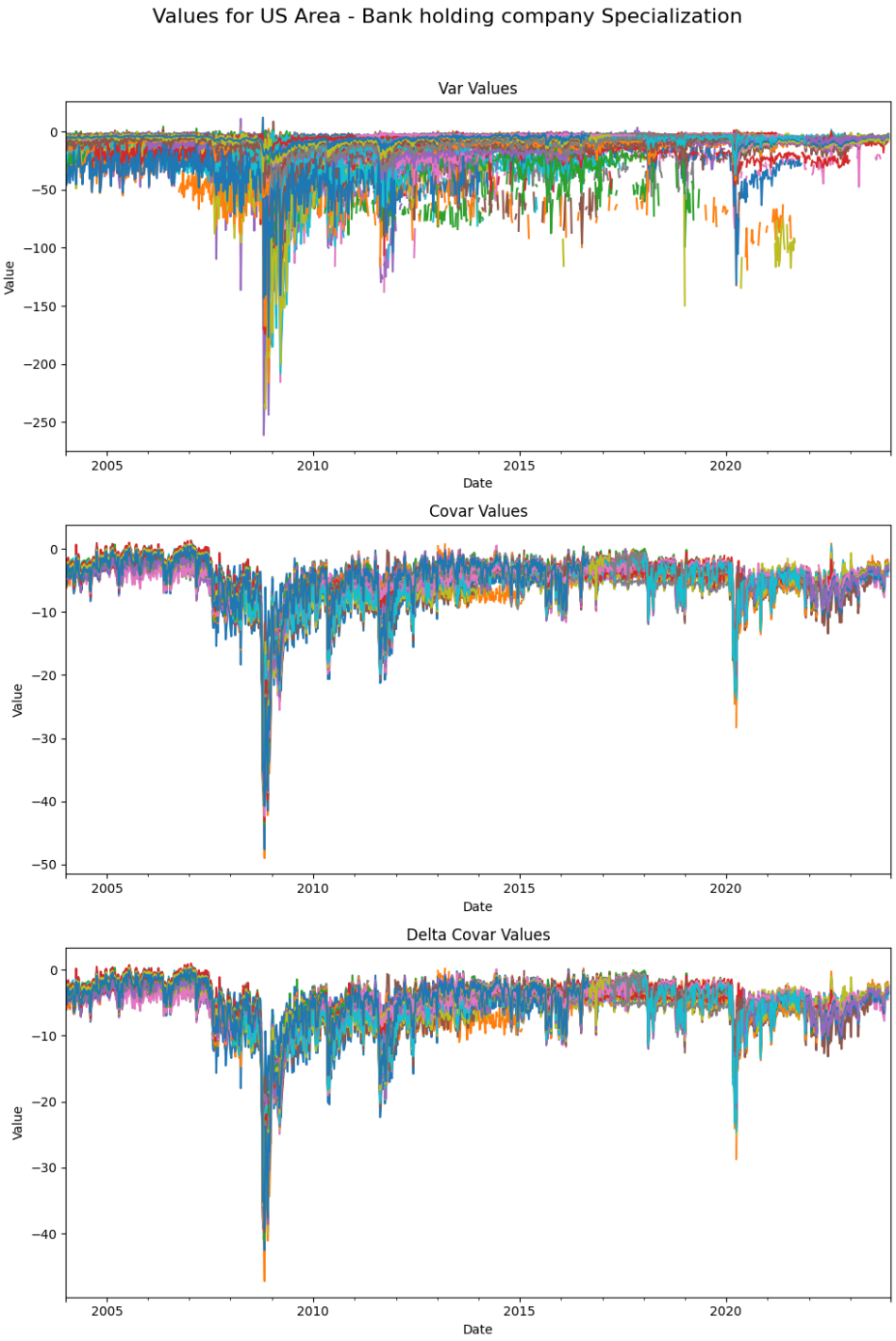


Figure 6.4: Var, CoVaR, ΔCoVaR for bank holding companies - US area (Inactive)

The graphs also reveal several other key insights:

1. While in crisis periods the individual risk (VaR) appears more 'clustered', the contribution to systemic risk (ΔCoVaR) seems to almost follow a homogeneous pattern. This seems to indicate that each institution's contribution to systemic risk is moving in a synchronized, or highly interconnected way, and that institutions are not just experiencing individual distress, but are also collectively contributing to a spiral of risk throughout the system.

2. There appears to be a weak link between VaR and ΔCoVaR, as some institutions with relatively low individual risk have a more pronounced contribution to systemic risk and vice versa. This distinction highlights that, while all institutions' risk measures increase during a crisis, their magnitudes differ.

3. Systemic risk is counter-cyclical, with the most significant increases in all risk measures occurring during crises when contemporaneous risk is high. This reinforces Adrian and Brunnermeier's (2011) [2] argument for employing, instead of contemporaneous risk measurements, a forward-looking measure (forward ΔCoVaR) based on stable characteristics to ensure more effective and non-procyclical regulation.

6.2 Crisis Periods

To analyze and compare the LQR and NNQR results, a more granulated level of detail is required. This section will comment on some of the most significant plots for specific sectors and for different geographical areas, but in this case only periods of financial crisis are considered. The periods of financial crisis considered range from slightly before the crisis to slightly

after, to see how the risk metrics change from before to after a financial crisis. As briefly mentioned in the previous section, the crisis affecting the data of the study that appear more relevant and visible in the graphs are:

- *Great Financial Crisis (GFC)*

  Caused by a subprime mortgage bubble and risky lending practices that led to the collapse of major financial institutions and a global recession.

  Period considered: 04/08/2006 – 25/12/2009

- *Covid-19 Pandemic*

  Lockdowns and supply chain disruptions which contracted global economic activity

  Period considered: 04/10/2019 – 29/12/2023

The plots in this section show some plot examples for the GFC and Covid periods. Although the plots only show commercial banks, the insights deduced from these examples appear to be valid for most of the banks' specializations in the dataset.

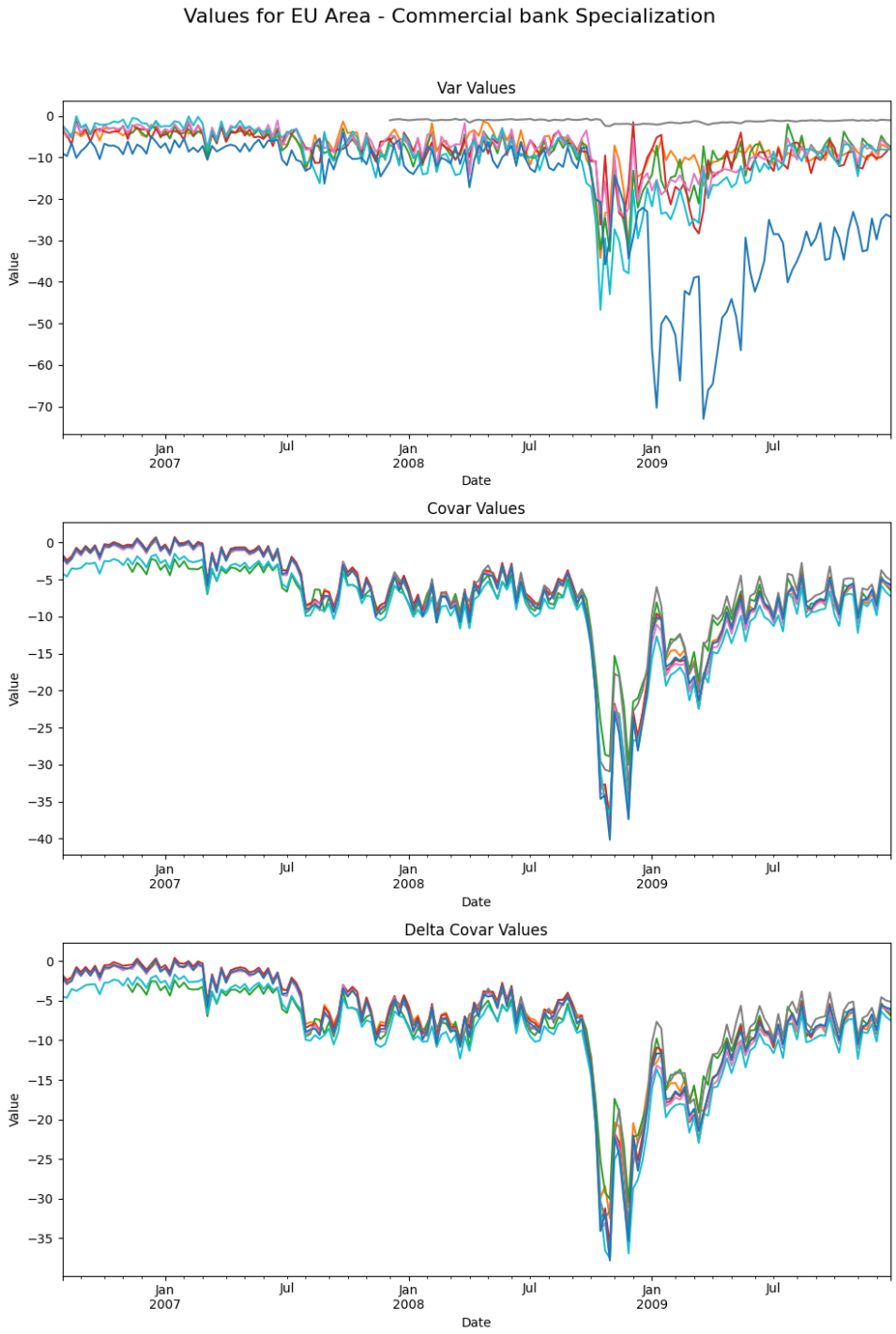


Figure 6.5: Var, CoVaR, ΔCoVaR (LQR) for commercial banks - Euro area (Merged)

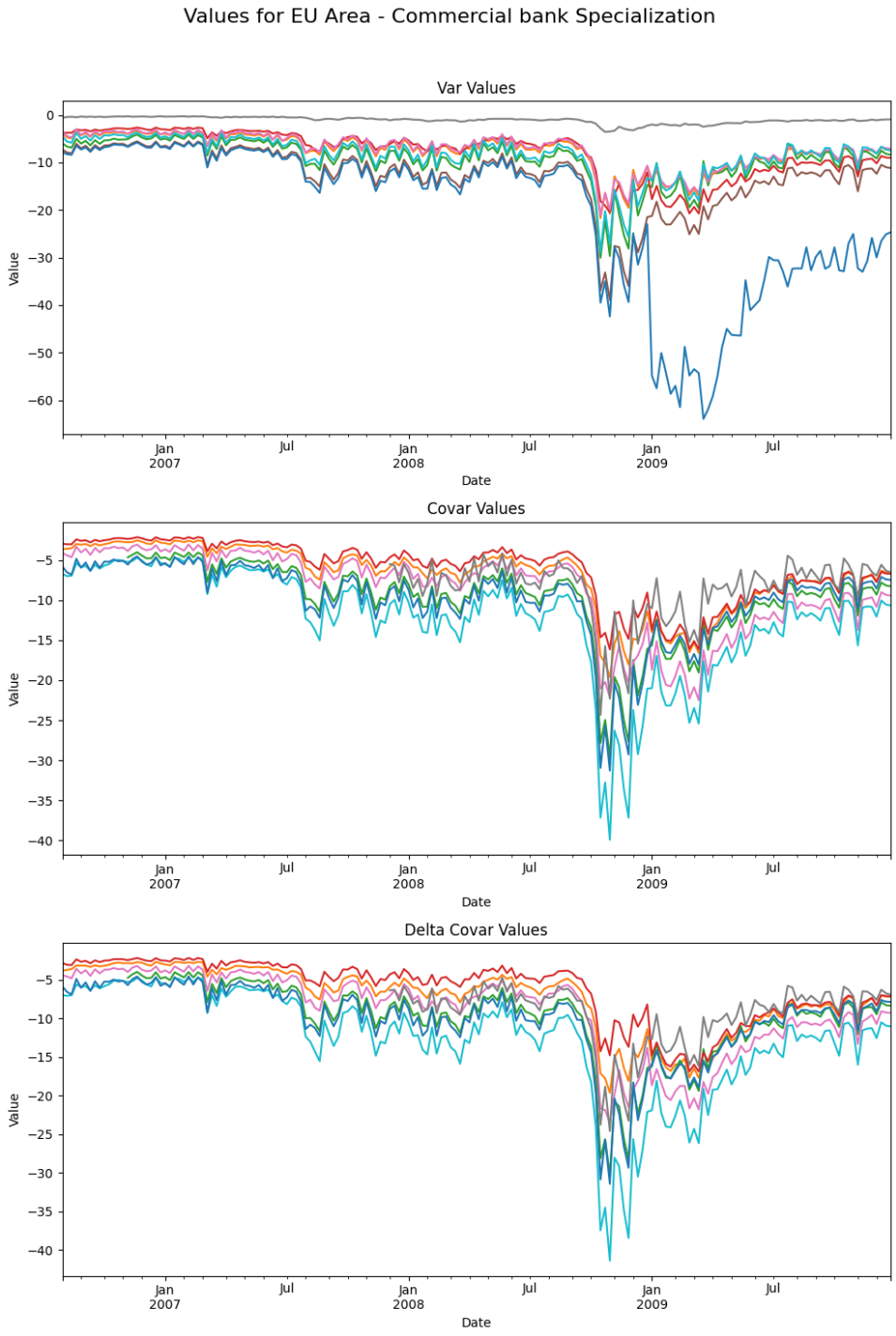


Figure 6.6: Var, CoVaR, ΔCoVaR (NNQR) for commercial banks - Euro area (Merged)

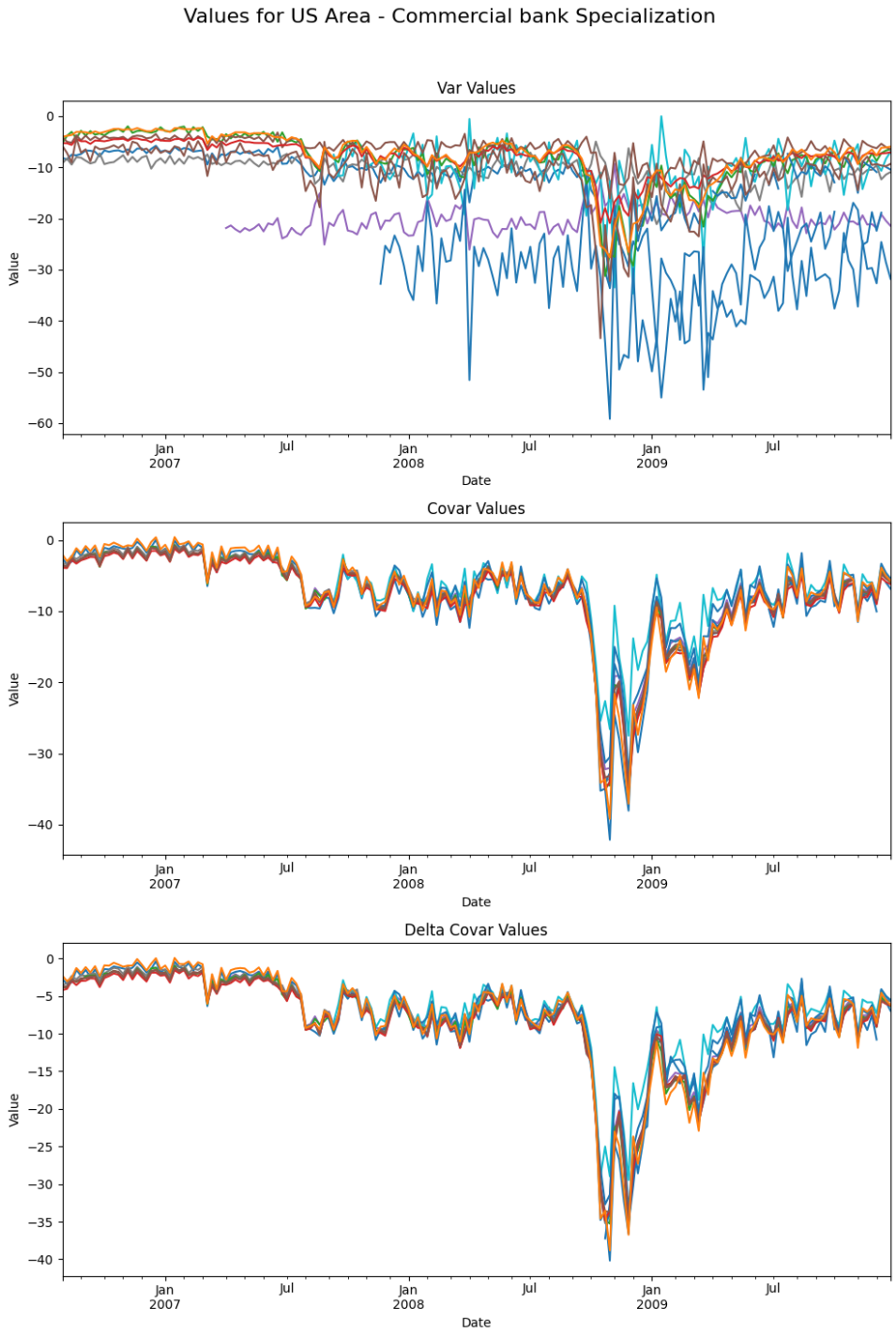


Figure 6.7: Var, CoVaR, ΔCoVaR (LQR) for bank holding companies - US area (Merged)

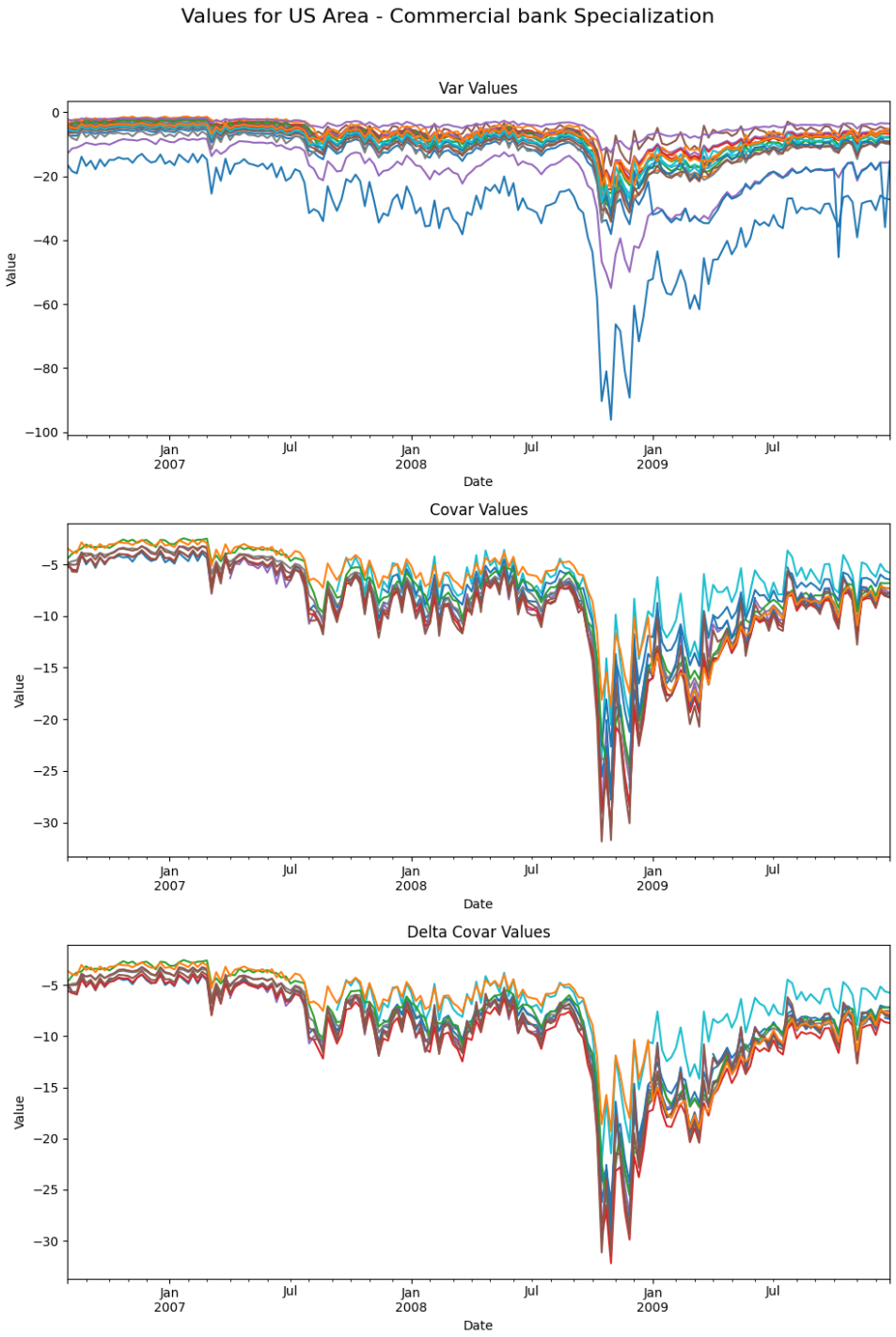


Figure 6.8: Var, CoVaR, ΔCoVaR (NNQR) for bank holding companies - US area (Merged)

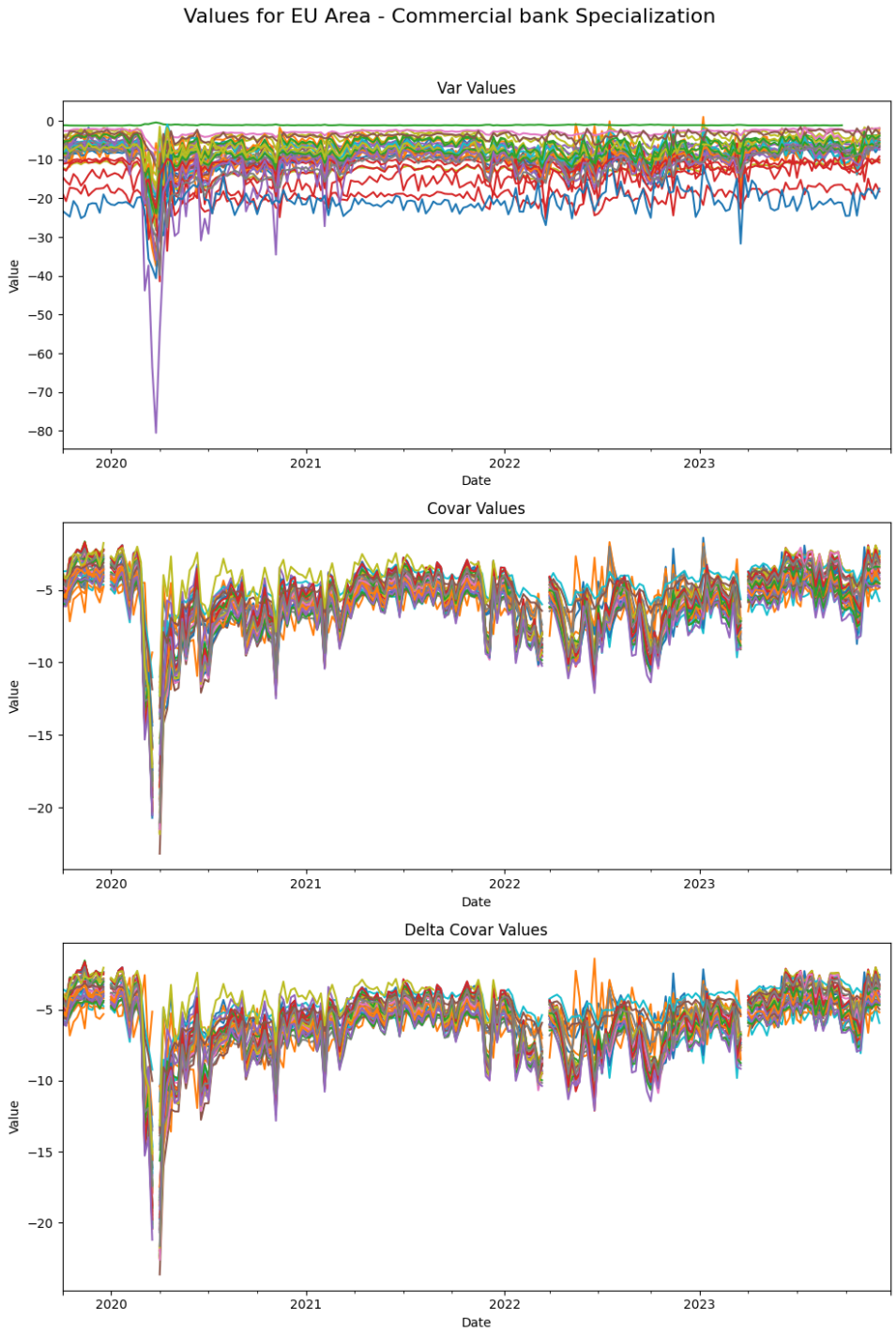


Figure 6.9: Var, CoVaR, ΔCoVaR (LQR) for bank holding companies - Euro area (Active)

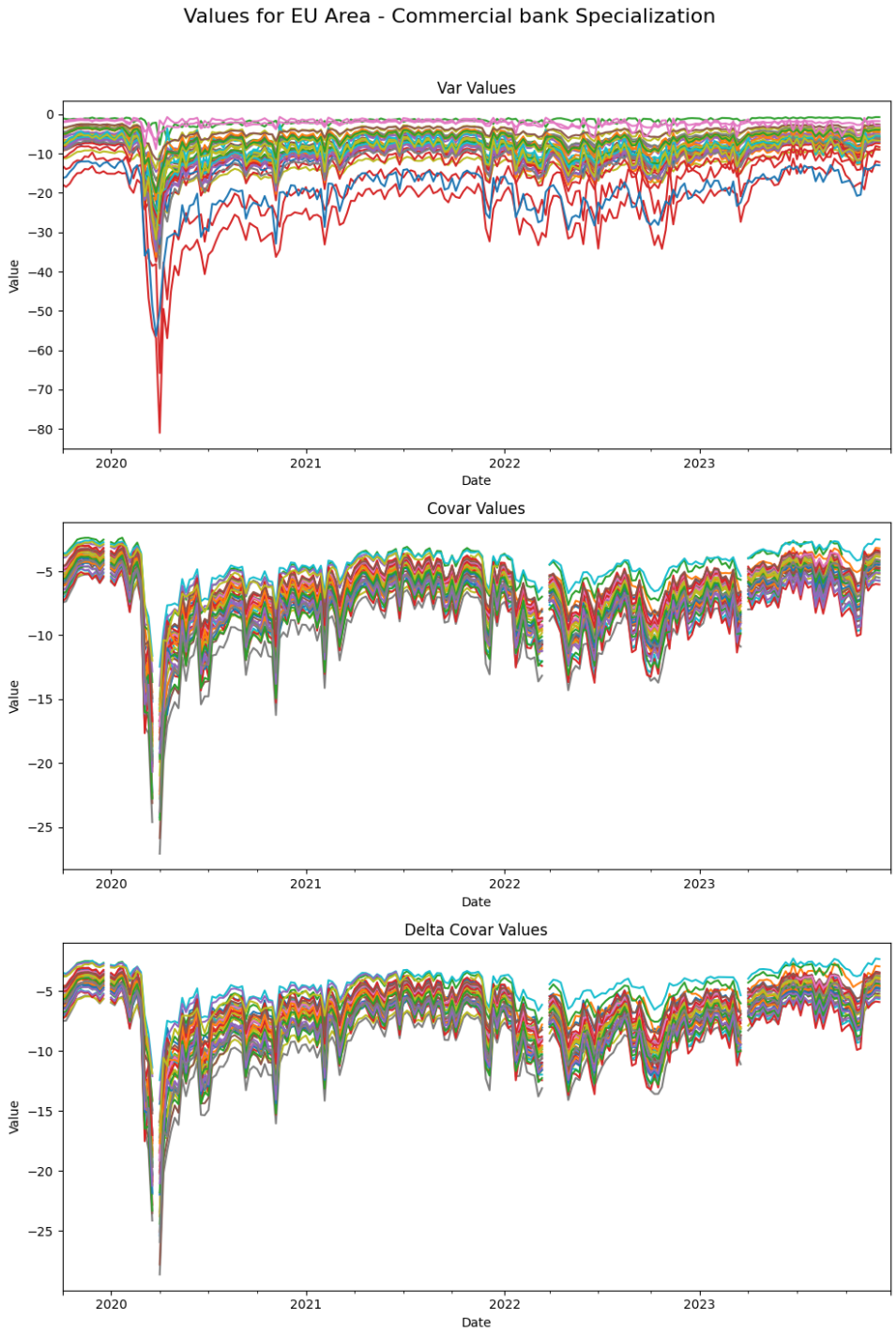


Figure 6.10: Var, CoVaR, ΔCoVaR (NNQR) for bank holding companies - Euro area (Active)

The plots 6.5 and 6.6 show VaR, CoVaR, and ΔCoVaR , from LQR and NNQR respectively, for commercial banks in the EU area during the GFC period. The plots 6.7 and 6.8 show the risk metrics from LQR and NNQR respectively, for commercial banks in the US area during the same crisis. The status group considered for GFC plots is the rest group merged with the inactive group. The reason behind choosing this status group only is that, since both the inactive and rest groups contain banks who have or had financial distress, it is very likely that most of these banks' distress occurred in the GFC, which caused extreme distress in the financial sector. To confirm this choice, the merged group of entities have the most interesting and significant plots for the analysis.

The VaR plots 6.5, 6.7, and 6.9 resulting from the LQR calculations are characterized by a more 'spiky' and volatile nature compared to those resulting from the NNQR in 6.6, 6.8, and 6.10. This difference arises from the LQR inability of capturing nonlinear dependencies and its attempt to fit a linear relationship in a nonlinear process. The linear model struggles to accurately represent the true distribution's tails, leading to sharp and discontinuous jumps in its VaR estimates. On the other hand, the NNQR's higher ability to map extreme events results in VaR estimates that more accurately reflect potential tail losses.

In contrast, the plots for CoVaR and ΔCoVaR 6.5, 6.7, and 6.9 from the LQR model appear smoother and more homogeneous. Unlike the 'clogged' plots from the previous section of this chapter, these period-specific plots should highlight more which entities have a higher impact of systemic risk, but instead the linear assumption prevents the model from detecting the nuances of each bank. The higher level of detail and the separation between the lines of the entities in the plots resulting from the NNQR shown in 6.6,

6.8, and 6.10 proves that NNQR can more accurately model the impactful and nonlinear spillover effects that were prevalent during the crisis.

In conclusion, the graphical analysis shows again, in accordance with the previous chapter, a slightly better performance of NNQR for systemic risk analysis as opposed to LQR.

Chapter 7

Forward-ΔCoVaR

As discussed in the methodology, Adrian and Brunnermeier (2011) [2] addressed the pro-cyclicality issues of contemporaneous risk measures by developing a forward-ΔCoVaR. This measure avoids the "procyclicality pitfall" (check the methodology chapter for further info) by estimating the relationship between the characteristics of the current firm and future spillover effects, proxied by ΔCoVaR_i. This section analyses the results of the panel regression computed on ΔCoVaR from both LQR and NNQR, using lagged variables (1 quarter lag, 1 year lag, 2 years lag). The regression aims at detecting the most significant variables that can predict the contribution of systemic risk 1 quarter, 1 year, or 2 years later. Separate calculations are made for the active group and the merged group (rest group and inactive group combined).

7.1 Correlation Matrix

Before proceeding with the forward ΔCoVaR, the correlation matrix of the variables and the consequent selection of variables were made. This step is relevant for panel regression because correlation between predictor variables

decreases the reliability of coefficient estimates, making it difficult to determine the individual impact of each correlated variable on the outcome.

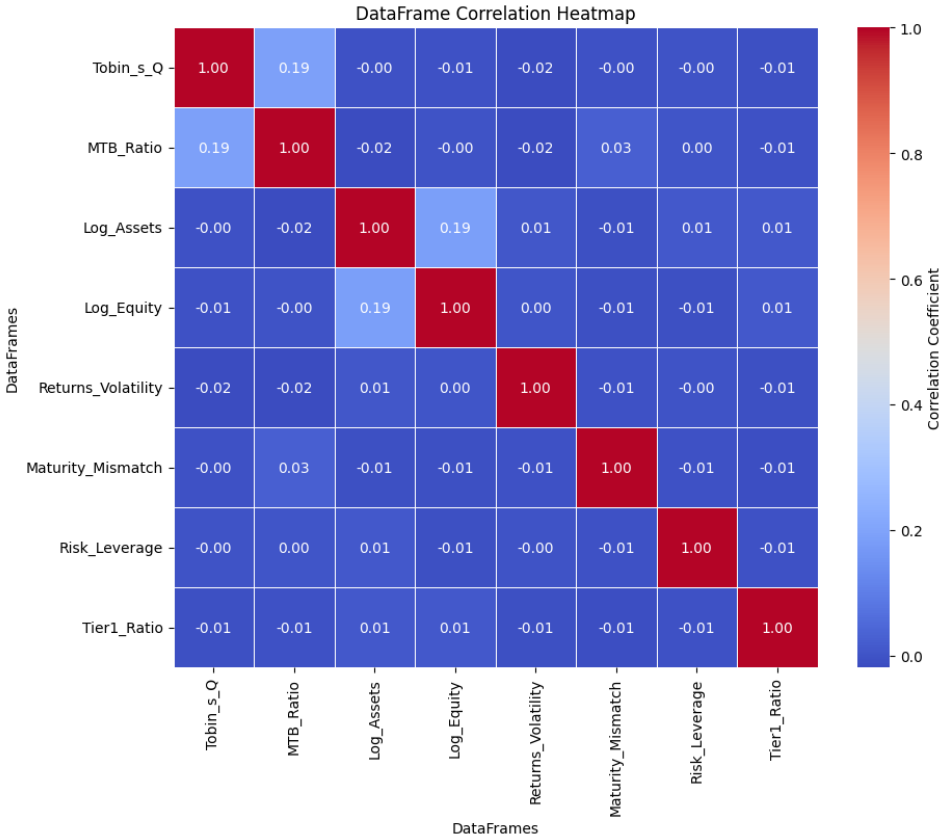


Figure 7.1: Panel regression variables correlation matrix (Active)

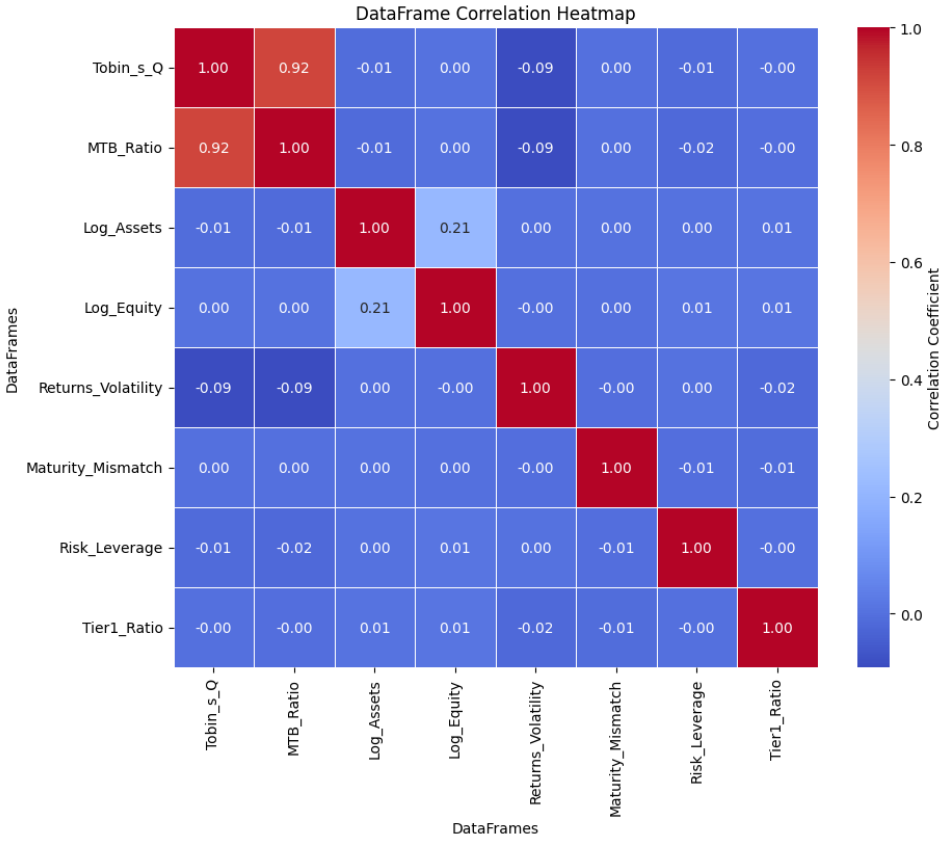


Figure 7.2: Panel regression variables correlation matrix (Merged)

Both the active group and the merged group show correlations between the Tobin's Q and the market-to-book ratio (MTB), and between the Log of Assets and Log of Equity. The first correlation is due to the fact that a high Tobin's Q implies that the market values the growth and future profitability of the entity, while a high MTB ratio indicates that the market values the entity more than its accounting-book values, and therefore both suggest a high positive sentiment (and vice versa in case of low values). However, in a distressed scenario such as insolvency, bankruptcy, liquidation, etc. investors become in fact more concerned with the underlying value of the firm's assets and liabilities, which are reflected in the book value. The Tobin's Q, which is usually also influenced by many factors such as future strategies, reputation, and others, during financial distress of an entity therefore converges to the MTB ratio. This explains well why the correlation between the Tobin's Q and the Market to Book ratio is extremely high for the merged status group, for which the Tobin's Q was consequently removed, and merely subtle for the active status group, for which both variables were kept.

The correlation between Log of Assets and Log of Equity is due to the basic accounting equation:

$Assets = Liabilities + Equity$

As a bank's total assets grow, both debt and equity must also increase. The positive correlation means that as a bank gets bigger (more assets), it is also increasing its equity base. However, since the correlation is not very significant, the two variables were kept in both the active group panel regressions and he merged group panel regressions.

Before running the panel regression, a static summary of the variables was made, revealing some interesting and yet expected findings. Return volatility and risk leverage of the merged group are much higher on average compared

to those of the active group, while Tier 1 ratios on average appear smaller in the merged group compared to the active group. The high volatility of returns and the high-risk leverage in the merged group confirm financial instability and distress, while Tier 1 points to low capital adequacy. Both these findings align with the characteristics of the merged group, as this group consists of financial distressed banks.

7.2 Panel regression

Before looking at the coefficients of the panel regressions, a few comments should be made on the performance characteristics of the panel regressions of the two versions of ΔCoVaR. First, the pseudo R-squared for the panel regressions on both LQR ΔCoVaR and NNQR ΔCoVaR for all lags are very low. The pseudo-R-squared, a metric described earlier, indicates the proportion of variance in a dependent variable that is explained by an independent variable(s) in a regression model. A low value indicates poor overall predictive power. This result can be attributed to the "balancing out" effect, a phenomenon known as heterogeneity. In the case of this study, while the fixed effects model accounts for average differences between entities, banks in different categories of specialization could have different behaviors compared to each other. For example, investment banks could be highly sensitive to certain market parameters, while others might not, or they could respond in different ways. Panel regression provides an average effect, but, as it is in this study, it cannot capture this varying influence between different specializations.

Another possible reason for the low R-squared is the presence of random noise from inaccuracies due to reporting errors or missing data reported as random numbers. The latter seems to be the case for some parts of Bloomberg's

data. Although the 'noisy' data were removed when detected, they are generally very difficult to find in a large data set. If not eliminated, these noisy values can substantially change the outcome of the predictions. The low values of pseudo R-squared can also be explained by the fact that the real relationship between the variables and the calculated ΔCoVaR is not linear and there is a more complex structure underneath the two that requires a more sophisticated model. Another possibility is that there are also other variables that could be included that might explain in a more statistically significant way the change in ΔCoVaRs.

7.2.1 Comparisons

Looking at the differences between the pseudo-R-squared of LQR ΔCoVaR and those of NNQR ΔCoVaR of the active and merged status groups, it emerges that the linear model performs better than the neural network for the active group. Neural networks, on the other hand, from a pseudo R-squared point of view, are much better for the merged status group, which includes entities with varying and potentially more complex financial statuses that cannot easily be mapped with a linear model. The data suggests that for these entities, particularly over longer time lags, the non-linear modeling capability of the neural network allows it to find better patterns, resulting in a better fit (higher R-squared and F-statistic).

An interesting finding is that for both models, the best predictive performance is achieved using the 1-year lagged variables. The 1-year lag is probably the most effective for both model types because it captures a crucial financial period because many financial metrics, such as annual earnings, fiscal year-end reports, and dividend payments, are reported on an annual basis. In terms of the business cycle, the 1-year delay captures its full effects,

demonstrating a more stable and complete picture of each entity's performance compared to a quarter-to-quarter view, which can be more volatile. In addition, many strategic decisions and their financial impacts, such as mergers, large-scale projects, and important capital expenditures, typically unfold over a year or more. Therefore, using data that reflect the full annual cycles provides the most informative features to predict the contribution of systemic risk, regardless of whether a LQR or NNQR model is used.

7.2.2 Coefficients

In financial terms, the interpretation of the coefficients for the panel regression of ΔCoVaR calculated with LQR or NNQR is complex, as their signs and significance vary not only between models, but also between different status groups. The following analysis will be centered on the 1-year lagged variables because, as mentioned above, they appear to have the best predictive power compared to the 1-quarter and 2-year lagged variables. Because Adrian and Brunnermeier's (2011) [2] methodology was followed, a more negative ΔCoVaR is associated with a greater systemic risk.

The first coefficients analyzed are from the panel regression on LQR ΔCoVaR of the active status group (with 1 year lagged variables) 7.1, which, as already mentioned, performed better than the panel regression on NNQR ΔCoVaR using the same variables.

The coefficients resulting from this panel regression might appear counter-intuitive at first glance. The positive coefficient of the log of assets and the negative coefficient of the log of equity indicate that the larger the size of the entity, the lower the systemic risk, and the higher the level of equity, the higher the systemic risk. In asset terms, this finding is in contradiction with the general idea that larger financial institutions, if they were to collapse,

Variable	Parameter	Std. Err.	T-stat	P-value	Lower CI	Upper CI
const	-140.07	1.4471	-96.799	0.0000	-142.91	-137.24
Tobin's Q	-0.878	1.0078	-0.8712	0.3836	-2.8532	1.0972
MTB Ratio	0.0501	0.0803	0.6239	0.5327	-0.1073	0.2075
Log Assets	7.1801	0.1893	37.939	0.0000	6.8092	7.5511
Log Equity	-3.7215	0.1553	-23.968	0.0000	-4.0258	-3.4171
Returns Volatility	-1.5316	0.3226	-4.7476	0.0000	-2.164	-0.8993
Maturity Mismatch	0.0272	0.0011	25.555	0.0000	0.0251	0.0293
Risk Leverage	-0.2292	0.0105	-21.741	0.0000	-0.2499	-0.2086
Tier1 Ratio	-0.2109	0.0099	-21.274	0.0000	-0.2304	-0.1915

Table 7.1: Panel regression for LQR ΔCoVaR - 1 year lagged variables (Active)

would cause widespread economic damage due to their key role in the financial system. For equity, generally, a higher equity is typically seen as a buffer against risk and not a possible factor causing it. A possible explanation for these results could be that firms with lower assets could fall into the category 'too small to care', resulting in less scrutiny from the regulatory agency. This factor can sometimes cause many risky behaviors or any other financial problem to go undetected for a long time. If many financial entities hold the same types of assets, a negative shock to those assets can affect multiple firms simultaneously. Consequently, while holding such a portfolio represents a sound diversification strategy for an individual firm, the collective adoption of that identical strategy results in a dangerous concentration of risk at the system level. If a widespread panic occurs, all these firms will try to sell the same assets at once, causing a crash that affects even healthy firms. In terms of capital structure, high equity reserves often enable institutions to scale up lending and trading exposures to riskier counter-parties. This expanding balance-sheet capacity weaves an intricate web of financial interdependencies that facilitates the rapid, system-wide propagation of distress. Additionally, high equity levels can foster risk-taking incentives. Rather than simply functioning as a protective buffer for regulators, substantial capital

reserves can serve as a cushion that emboldens management to engage in speculative investments to boost returns. Combined with high interconnectedness, the failure of these aggressive bets can ignite a widespread systemic contagion.

Both lagged Tobin's Q and the lagged Market-to-Book ratio are statistically insignificant in the panel Linear Quantile Regression (LQR) of ΔCoVaR for the active group. This suggests that an institution's market valuation or growth expectations in a given year do not meaningfully influence its systemic risk contribution (ΔCoVaR) in the subsequent year. Maturity mismatch—defined as funding long-term assets with short-term liabilities—unexpectedly yields a positive coefficient, indicating that higher mismatch is associated with lower subsequent systemic risk. A plausible explanation is regulatory scrutiny: institutions with pronounced maturity mismatches face heightened oversight, prompting them to actively derisk and rebalance their portfolios over the following year. Conversely, firms exhibiting initially low mismatches may attract less supervisory attention, allowing unmonitored liquidity vulnerabilities to accumulate undetected over the same period.

Return volatility and risk leverage coefficients follow a more conventional pattern. Their negative values indicate that entities with more volatile returns and entities that were more leveraged a year before are more likely to have a higher systemic risk in the following year. The tier 1 ratio coefficient for this ΔCoVaR's panel regression is surprisingly negative, indicating that a higher ratio increases systemic risk. The Tier 1 ratio is a bank's Tier 1 capital (primarily equity) divided by its risk-weighted assets (loans and investments). A higher ratio means that the bank is legally required to have more equity relative to its lending and investment activities, so when regu-

lators set a higher capital requirement, they are explicitly telling banks to fund their operations with a greater proportion of equity. Forcing banks to rely on this more expensive source of funding increases their overall cost of capital, which in turn leads to a reduction in profitability and an increase in the cost of borrowing for their customers. This dynamic can then lead to, as banks reduce lending, a credit crunch that stifles economic growth and the shift of some lending activities to the less-regulated shadow banking system (e.g., hedge funds and money market funds) a hidden and less-monitored area of finance, where a rapid build-up of leverage can occur. The lack of transparency and oversight in the shadow banking system makes it difficult to manage risk and increases the likelihood of a crisis that can spill over into the traditional banking system.

The other coefficients analyzed are from the panel regression on NNQR ΔCoVaR of the merged status group (active rest and inactive groups together), with 1 year lagged variables, 7.2. As already mentioned, for this group the NNQR performed better than the panel regression on LQR ΔCoVaR using the same variables.

Variable	**Parameter**	**Std. Err.**	**T-stat**	**P-value**	**Lower CI**	**Upper CI**
const	-113.14	1.718	-65.851	0.0000	-116.5	-109.77
MTB Ratio	-1.3866	0.1185	-11.699	0.0000	-1.6189	-1.1543
Log Assets	-1.574	0.2591	-6.0755	0.0000	-2.0818	-1.0662
Log Equity	2.7426	0.1479	18.542	0.0000	2.4527	3.0325
Returns Volatility	-0.9647	0.8511	-1.1335	0.257	-2.6329	0.7034
Maturity Mismatch	0.0076	0.0045	1.7108	0.0871	-0.0011	0.0164
Risk Leverage	-0.0474	0.0032	-14.617	0.0000	-0.0537	-0.041
Tier1 Ratio	-0.0653	0.015	-4.3524	0.0000	-0.0947	-0.0359

Table 7.2: Panel regression for NNQR ΔCoVaR - 1 year lagged variables (Merged)

The values of these coefficients seem to be more in line with general financial

assumptions. The log of assets has a negative coefficient that indicates that higher values lead to lower ΔCoVaR and, therefore, to a higher systemic risk. This aligns with the 'too big to fail' theory that larger firms have a higher chance of fueling a chain of systemic risk reactions across all the financial institutions. The positive coefficient for the log of equity, signaling that higher equity is associated with lower risk, also follows the traditional idea that a strong capital buffer provides a cushion against financial distress. In terms of market-to-book ratio, while a higher ratio often reflects expectations of future growth and profitability, in the contest of higher values leading to higher systemic risk, as the negative coefficient indicates, it could signal that often those expectations are a consequence of a speculative bubble or any other reason leading to an inflated valuation.

As in the case of the linear ΔCoVaR discussed earlier, a higher risk leverage at a given time leads to a higher systemic risk the following year. This is a classic finding in financial risk analysis, as high leverage can amplify losses and increase the risk of bankruptcy and contagion. The tier 1 ratio coefficient is also negative in this case just as in the case previously discussed. Although the coefficient is lower for the merged status group ΔCoVaR calculated with NNQR, its interpretation remains the same. For maturity mismatch and volatility of returns, they both seem statistically not significant.

7.3 Conclusion

The interpretation of the coefficients of the previous chapters shows that the superior performance of NNQR ΔCoVaR for the merged status groups indicates an inherently nonlinear, complex, and interconnected nature of financial risk in the precrisis era. In fact, in the precrisis regulatory environment, before the implementation of frameworks such as Basel III, capital

requirements were less demanding and systemic risk was not as central a focus for regulators. Financial institutions operated under rules that did not fully account for interconnectedness and potential for contagion. As a result, when these firms were approaching distress, their financial metrics followed what today would be considered a predictable pattern. A lack of capital (low equity) was a direct indicator of vulnerability, and larger size truly meant "too big to fail", as evidenced by the widespread economic damage that followed their collapse. Because the entities in the merged grouped manly faced financial struggles and consequently failed or faced merges after the great financial crisis, most of their financial data stop after that period or a few years after, and therefore the ΔCoVaR analysis almost entirely focus on that time frame.

The performance of the panel regression on NNQR ΔCoVaR compared to the one on LQR ΔCoVaR suggests that although the risk profile of the merged group was traditional, the interactions between these risk factors were not simple or linear. Instead, these factors likely exhibited a complex pattern that culminated in bank failure (or similar), a pattern to which a nonlinear model is better suited to detect.

For the active status group, it is interesting how the panel regression on the LQR ΔCoVaR shows very unconventional patterns. As mentioned above, the reason for this is probably related to the fact that after the great financial crisis, regulators focused heavily on capital buffers and liquidity. This factor led to increased monitoring of the "too big to fail" institutions, which can in turn lead to new subtle forms of systemic risk such as the 'moral hazard' of a strong balance sheet, where management could be encouraged to pursue riskier, higher-yield strategies as they feel protected by their high capital cushion. There is also the " too small to care" effect, which allows

smaller, less-capitalized institutions to operate under the radar of regulators and engage in risky behaviors that, while not individually systemically important, can, collectively, pose a significant threat.

However, determining why the panel regression on the LQR ΔCoVaR works better than the NNQR version for the active group is a more challenging task. There could be many explanations for this.

First, the superior performance of the LQR ΔCoVaR model could stem from the fact that a panel regression, by its nature, is better suited for the simpler linear relationships captured by the LQR ΔCoVaR measure and therefore a simpler systemic risk measure, derived from a linear model, is better explained by a simple linear panel regression. In contrast, the NNQR ΔCoVaR measure, which is inherently more complex, may have underlying non-linear patterns that a standard panel regression model struggles to fully capture or explain, even if it is a technically more advanced and accurate risk measure. NNQR ΔCoVaR works for the merged group (inactive banks and distressed active entities) though , which appears to be contradicting the previous statement. In reality, this occurs because distressed and inactive institutions have severe spikes in balance-sheet deterioration, allowing the linear panel regression to still easily map these large ΔCoVaR spikes to the underlying firm variables, while in the case of active companies the links are more subtle.

The second possible reason for a better performance of LQR on ΔCoVaR for active companies is that the chosen variables do not fully capture the more complex ΔCoVaR calculated with NNQR and therefore a different, perhaps broader choice of variables could lead to different results.

Therefore, a combination of nonlinear model choice for computing the forward-ΔCoVaR and a broader, perhaps different, selection of variables will be more likely to achieve better results.

Chapter 8

Final Remarks

The seminal work by Adrian and Brunnermeier (2011) [2] on ΔCoVaR, a metric designed to quantify the contribution of systemic risk of banking entities and predict future spreads of systemic risk, remains a powerful but underutilized tool in its full potential. Despite the proliferation of advanced data science and machine learning techniques, the widespread application and integration of this model into modern financial practice has been limited to the linear case. Although the broader financial industry has begun to take advantage of the exponential growth of data in the digital age, banking regulatory bodies have historically been slower to adopt such innovations. However, a small shift is happening. The European Central Bank (ECB), for example, is actively investigating the potential of artificial intelligence (AI) to improve the efficiency and quality of its supervision work, creating a new environment for re-evaluation and modernization of existing models [18].

Recognizing the potential ΔCoVaR, this study aims to address the existing lack of advanced techniques in regulatory frameworks and in the literature, by showing that ΔCoVaR can be significantly improved by applying modern

machine learning techniques. Specifically, the core objective of this research was to improve the initial Adrian-Brunnermeier's model by comparing they original LQR model with its neural network quantile regression (NNQR) version. Furthermore, the study also applies a panel regression analysis using a set of lagged institutional variables following Adrian-Brunnermeieron's methodology, to provide a comparison of the two types of ΔCoVaRs in a systemic risk prediction context.

The main conclusion of this research is that the NNQR-based ΔCoVaR model demonstrated a marginal but meaningful improvement over the original Adrian and Brunnermeier LQR-based ΔCoVaR. Although these findings validate the efficacy of integrating advanced nonlinear techniques into econometric models for systemic risk, there is still considerable room for further research.

A significant limitation of the study lies in its reliance on US market state variables for the calculations of VaR and CoVaR, even for European institutions. Although the influence of the US market on the global financial system is undeniable, this approach risks neglecting the unique characteristics of the European financial landscape. This oversimplification could have led to a loss of relevant information that could have been relevant for a more accurate assessment of systemic risk within the European context. Furthermore, the data preparation methodology also presents an opportunity for refinement. The use of simple linear interpolation for missing financial data points, while straightforward, may not capture the true dynamics of the financial time series. More robust imputation techniques, such as Multiple Imputation by Chained Equations (MICE), could provide a more statistically sound and representative dataset, thereby also improving the reliability of the model's output. Furthermore, incorporating asymmetries into the model by com-

puting CoVaR using separate coefficients for negative and nonnegative bank returns, as done by López-Espinosa et al. (2012) [15], could produce significant improvements because it would allow the model to capture the system sensitivity to bank returns during both periods of loss and periods of gain.

The reliance on traditional panel regression also imposes restrictions on the model's predictive power. As a linear method, it may not fully capture the complex, non-linear interactions inherent in systemic risk. Future research could pivot to more flexible and powerful machine learning algorithms, such as gradient-boosting models like XGBoost, to better model these intricate relationships. Although these "black-box" models present challenges for interpretability, techniques such as Individual Conditional Expectation (ICE) plots can be employed to gain valuable insight into how the financial characteristics of single institutions influence systemic risk contribution (through ΔCoVaR). A more sophisticated data aggregation strategy for ΔCoVaR in the panel regression could also be beneficial. The current practice of summing the data for quarterly conversion and handling missing values simply by inserting NaNs for the entire quarter introduces data loss and potential bias. Applying a more advanced imputation technique like MICE to address these missing values would result in a more complete dataset for the panel regression.

Ultimately, the general difficulty of implementing a sophisticated machine learning model for systemic risk analysis is increased by a number of factors. First, the computational power required to train and optimize complex models, particularly deep neural networks, can be a substantial barrier. Second, access to data on interconnections for research, such as interbank assets and liabilities, is a challenge: much of this information is proprietary or its use is restricted to regulatory bodies. This lack of some important data limits

the potential for significant methodological improvements and exploration of new variables and geographical areas. Therefore, an import takeaway from this research is that meaningful progress in systemic risk prediction requires the concerted effort of multiple stakeholders. In particular, there is an increasing need for regulatory bodies to promote responsible data sharing with the academic community and champion the use of AI, thereby bridging the gap between theoretical advances and their impactful practical applications.

Bibliography

[1] Pawel Smaga. The concept of systemic risk. *Systemic Risk Centre Special Paper, The London School of Economics and Political Science*, 5, 08 2014.

[2] Tobias Adrian and Markus K. Brunnermeier. Covar. *National Bureau of Economic Research Working Paper Series*, No. 17454, October 2011 2011.

[3] James W. Taylor. A quantile regression neural network approach to estimating the conditional density of multiperiod returns. *Journal of Forecasting*, 19(4):299–311, 2000.

[4] Franklin Allen and Elena Carletti. What is systemic risk? *Journal of Money, Credit and Banking*, 45(s1):121–127, 2013.

[5] Philippe Jorion. *Value at risk 3E.* Tata McGraw-Hill Education, December 2009.

[6] Paul H. Kupiec. *Stress Testing in a Value at Risk Framework*, page 76–99. Cambridge University Press, 2002.

[7] Roger Koenker and Gilbert Bassett. Regression quantiles. *Econometrica*, 46(1):33–50, 1978.

[8] Alex J. Cannon. Quantile regression neural networks: Implementation in r and application to precipitation downscaling. *Computers & Geosciences*, 37(9):1277–1284, 2011.

[9] Georg Keilbar and Weining Wang. Modelling systemic risk using neural network quantile regression. *Empirical Economics*, 62(1):93–118, Jan 2022.

[10] Gaurav Kumar, Molla Ramizur Rahman, Abhinav Rajverma, and Arun Kumar Misra. Predicting systemic risk of banks: a machine learning approach. *Journal of Modelling in Management*, 19(2):441–469, 07 2023.

[11] Shan Wu, Mu Tong, Zhongyi Yang, and Tianyi Zhang. Interconnectedness, systemic risk, and the influencing factors: Some evidence from china's financial institutions. *Physica A: Statistical Mechanics and its Applications*, 569:125765, 2021.

[12] Vicente Balmaseda, María Coronado, and Gonzalo de Cadenas-Santiago. Predicting systemic risk in financial systems using deep graph learning. *Intelligent Systems with Applications*, 19:200240, 2023.

[13] Gregor N.F. Weiß, Denefa Bostandzic, and Sascha Neumann. What factors drive systemic risk during international financial crises? *Journal of Banking & Finance*, 41:78–96, 2014.

[14] João Gabriel de Moraes Souza, Daniel Tavares de Castro, Yaohao Peng, and Ivan Ricardo Gartner. A machine learning-based analysis on the causality of financial stress in banking institutions. *Computational Economics*, 64(3):1857–1890, Sep 2024.

[15] Germán López-Espinosa, Antonio Moreno, Antonio Rubia, and Laura Valderrama. Short-term wholesale funding and systemic risk: A global covar approach. *Journal of Banking & Finance*, 36(12):3150–3162, 2012. Systemic risk, Basel III, global financial stability and regulation.

[16] Carlo Bellavite Pellegrini, Peter Cincinelli, Michele Meoli, and Giovanni Urga. The contribution of (shadow) banks and real estate to systemic risk in china. *Journal of Financial Stability*, 60:101018, 2022.

[17] Simone Varotto and Lei Zhao. Systemic risk and bank size. *Journal of International Money and Finance*, 82:45–70, 2018.

[18] Elizabeth McCaul. From data to decisions: Ai and supervision. `https://www.bankingsupervision.europa.eu/press/interviews/date/2024/html/ssm.in240226~c6f7fc9251.en.html`, 2024.