Konstantin Kundegraber, Bsc

Quantifying Regulatory Impact with NLP: How Enforcement Actions and their Legal Notices Affect Bank Stock Prices

Master's Thesis

Supervisor
Roman Kern, PhD

Institute of Interactive Systems and Data Science
Head: Frank Kappe, Dipl.-Ing. Dr.techn.

Second Supervisor
Andrea Schertler, Dr.

Institute for Banking and Finance
Head: Stefan Palan, Dr. Mag.

Graz, January 2024

Abstract

This thesis examines how the announcement of enforcement actions by the Office of the Comptroller of the Currency (OCC), Federal Deposit Insurance Corporation (FDIC) or Federal Reserve System (Fed) affects the stock prices of enforced U.S. banks. Using an event study approach, cumulative abnormal returns are calculated for different event windows. These returns are then tested for significance and further examined to capture predictive differences. In addition, Natural Language Processing (NLP) techniques, including large language models developed by OpenAI and Google, are employed to evaluate the risk associated with legal notices regarding enforcement actions taken against the bank. The results indicate negative cumulative abnormal returns for the enforced banks, but do not provide further insight into the variance of these returns. The thesis concludes that enforcement actions have a significant impact on the banks' stock prices.

Contents

List of Figures

List of Tables

1. Introduction

The 2008 Global Financial Crisis served as a wake-up call, revealing flaws in financial stability that led to a surge in enforcement actions against U.S. banking organizations. These actions are a direct response to the identification of unsafe illegal activities and allow regulators to impose sanctions and enforce corrective steps. By enforcing these measures, banks are pressured to adopt safe practices or modify their behavior to ensure stability and soundness. On the other hand, banks frequently consider enforcement actions to be challenging due to the direct and indirect costs they impose. Banks are responsible for allocating resources to address these issues. They may also be required to pay fines, provide financial compensation or offer refunds to those individuals who have been impacted. Additionally, it is important to note that sanctioned banks may suffer substantial harm to their reputation as a result of the public disclosure of these actions, which could further affect investors' stance towards the bank.

The objective of this thesis is to examine the impact of enforcement actions announced by the OCC, FDIC or Fed on the stock price of the corresponding U.S. bank. To do this, an event study was conducted to calculate cumulative abnormal returns and a statistical model was developed that attempted to explain the variation in the returns. This thesis, in contrast to existing approaches, utilized NLP operations. To achieve this, language models developed by OpenAI and Google were used to classify legal texts related to enforcement actions taken against the bank. Classified correctly, this enhances the ability to distinguish between different enforcement actions and explore potential opportunities for using language models in legal notice classification. Although stock-related research often involves simulating and assessing the performance of developed models, this thesis primarily focuses on theoretical aspects and does not cover practical implementations as this would be beyond the scope of the study.

The first part provides a detailed description of U.S. bank supervision, including the supervisory agencies involved and the potential types of enforcement they employ. This is followed by a brief introduction to returns and related literature. It provides insights into related work that has examined enforcement actions and event studies with the aim of providing a comprehensive overview of the topic. The thesis then presents the hypotheses, followed by an explanation of the materials and methods used. This section includes the presentation of the initial data and details of the processing techniques and APIs used. The event study methodology is also explained in detail, along with the hypothesis testing and models used for the cross-sectional analysis. The paper then presents the findings in the discussion section, which also includes a sensitivity analysis of the large language models used. Finally, the paper finishes with a conclusion.

2. Background & Related Work

2.1. U.S. Bank Supervision

The U.S. banking system is subject to strict supervision and regulation to ensure its stability and to protect the interests of depositors and the economy as a whole. This section examines the roles and responsibilities of the Office of the Comptroller of the Currency (OCC), the Federal Deposit Insurance Corporation (FDIC) and the Federal Reserve System (Fed) in supervising U.S. banks and the various types of enforcement actions that may be taken.

2.1.1. Supervising Agencies

The supervision is carried out by several key agencies, including the OCC, FDIC and Fed. Together, these agencies play a critical role in supervising and regulating banks, promoting safe banking practices and maintaining the integrity of the financial system.

Office of the Comptroller of the Currency (OCC) The OCC was established by President Abraham Lincoln in 1863 to organize and administer the banking and monetary system of the United States, creating a system of nationally chartered banks. Today, the OCC supervises more than 1,063 banks, including national banks, federal savings associations and federal branches and associations of foreign banks. These regulated institutions hold about $15 trillion in assets, representing 66 percent of all U.S. commercial banking assets. The OCC plays a crucial role in the regulation and supervision of these national banks and federal savings associations. It does this by providing banking rules and regulations, supplying legal interpretations and evaluating applications for new bank charters or branches. In addition, the OCC regularly visits and examines the banks it supervises to ensure compliance with laws and regulations. In cases where banks fail to comply or engage in unsafe practices, the OCC has the authority to impose corrective action. Finally, the OCC focuses on protecting consumers by ensuring fair access, equal treatment and adherence to consumer banking regulations (OCC, 2023a).

Federal Deposit Insurance Corporation (FDIC) In response to the extensive bank failures that occurred during the Great Depression, the FDIC was established in 1933. It plays a crucial role in protecting depositors' funds by offering deposit insurance to ensure that depositors do not lose their funds due to bank failure. As an independent agency of the federal government, it is tasked with the objective of maintaining stability and public trust in the country's financial system. In the event of a bank failure, the FDIC responds immediately to protect insured depositors by selling the deposits of the failed institution to another institution, ensuring a seamless transition for customers. The FDIC is responsible for supervising and examining banks to ensure their operational safety and soundness and compliance with consumer protection laws, as well as administering receiverships. It is funded by premiums paid by banks and savings associations, provides insurance for trillions of dollars of deposits in banks in the United States (FDIC, 2023a).

Federal Reserve System (Fed) The Federal Reserve Act was signed into law by President Woodrow Wilson in 1913, establishing the Fed. The Fed is composed of a Board of Governors and 12 regional Federal Reserve Banks spread across the United States. As the central bank of the United States, the Federal Reserve System plays a critical part in promoting the efficient functioning of the U.S. economy and serving the public interest. To achieve these goals, it performs five key functions. First, it carries out monetary policy in order to uphold the goals of achieving optimal employment levels, ensuring price stability and promoting moderate long-term interest rates within the economy. Second, it works to maintain the stability of the financial system by actively monitoring and intervening in domestic and international markets to minimize systemic risks. Third, it enhances the safety and stability of individual financial institutions while evaluating their overall impact on the broader financial system. Fourth, the Federal Reserve promotes the safety and efficiency of payment and settlement systems, facilitating U.S. dollar transactions and payments for the banking industry and the government. Fifth, it focuses on consumer protection and community development through supervising, evaluating, researching and analyzing emerging consumer concerns, as well as enforcing consumer laws and regulations (Fed, 2023a).

2.1.2. Enforcement Types

Enforcement actions are taken to address various violations such as breaches of fiduciary duty, unsafe practices and violations of laws or regulations. Understanding the various types of enforcement actions is crucial as it enables a comprehensive comprehension of the potential consequences and risks associated with these actions. Knowledge of the different types of enforcement actions can also aid in evaluating the financial impact or reputational harm that a bank may face.

Section 19 Letters Section 19 letters are a means of communicating with persons who have been convicted of certain offences or who have entered into a pretrial diversion program in relation to dishonesty, breach of trust or money laundering. Section 19 letters are official notifications that inform individuals they are prohibited from participating in the activities of insured depository institutions, their holding companies, or credit unions without obtaining prior regulatory approval or judicial authorization (Fed, 2023b).

Cease and Desist Orders A cease and desist order is a legal directive issued that requires a bank or a person associated with a bank to immediately cease and desist from engaging in an unsafe or unsound practice or violation. It may also require them to take the necessary steps to remedy any problems arising from such violations or practices and serves as a means of enforcing compliance (OCC, 2023b).

Prohibitions Prohibitions are legal measures that restrict an Institutional Affiliate Party (IAP) from participating in certain activities or roles. For example, an order may be made prohibiting an individual from participating in any way in the affairs of an insured depository institution. Similarly, individuals who have been indicted for certain crimes may be temporarily suspended or prohibited from participating in the affairs of an insured depository institution (OCC, 2023b).

Civil Money Penalty Banking organizations and IAPs may be subject to a civil money penalty if they participate in risky banking practices, breach regulations or neglect to follow directives from the banking regulator. In such cases, the regulators can issue orders requiring either the bank or an individual to pay fines (OCC, 2023b).

Written Agreement A written agreement signed by the banking regulator and the board of directors on behalf of a bank serves as a legally binding document outlining the terms and conditions agreed to by all parties (OCC, 2023b).

Deposit Insurance Termination A deposit insurance termination can be initiated if a banking organization is considered to be in an unsafe or unsound condition or if it has been involved in unsafe banking practices or breaches of the law. In such cases, the institution is required to notify all depositors before deposit insurance is terminated (FDIC, 2023b).

Prompt Corrective Action Insured banks must comply with certain restrictions and measures based on their capital category. These restrictions become mandatory once the bank is notified or becomes aware of its capital category. In addition, the regulator has the authority to impose restrictions and actions through the issuance of a prompt corrective action directive (OCC, 2023b).

2.2. Enforcement Actions

This section provides a comprehensive examination of the consequences of enforcement actions against banks. Understanding the impact of these enforcement actions is critical as regulators around the world continue to tighten their grip on the financial industry. Therefore, examining the consequences of these actions can further provide valuable insights into the effectiveness of these measures and their ability to influence a bank's behavior.

There is a large body of research on the impact of enforcement actions. According to Delis et al. (2017), enforcement actions lead to a reduction in risk-weighted assets and in the share of non-performing loans of the enforced banks. Interestingly, the earlier an action is taken, the faster banks take steps to improve their financial soundness. Conversely, the longer it takes, the higher the risk of the bank getting into serious difficulties. Moreover, banks subject to enforcement action are less likely to expand lending and the enforcement action is responsible for a significant decline in gross assets (Peek and Rosengren, 1995).

Brous and Leggett (1996) examined bank investors' perceptions of enforcement actions. They studied 62 cases against 61 banks and found that the average abnormal return on the day of announcement was -4.5 percent and that 82 percent of the banks studied had negative abnormal returns. They also found that the effect was larger for more diversified banks of considerable size as well as for parent companies than for their subsidiaries. Jordan et al. (1999) support these findings as they also observed large negative abnormal stock returns for banks on announcement days.

The consequences of enforcement actions are not limited to banks and their stakeholders but can also affect the real economy. Danisewicz et al. (2018) found that severe enforcement measures can have effects that go well beyond a bank's own operations. Enforcement measures significantly impact both real per capita income growth and the unemployment rate. However, it should be noted that this is not due to enforcement actions against individuals, which do not usually trigger adjustments in bank behavior.

2.3. Studying Stock Price Responses

The study of stock returns, which are basically the change in the price of a stock, has been the subject of research for a long time. Early research included a study of the distribution of stock returns. It was found that these distributions had „fat-tailed“ characteristics relative to a normal distribution, while the standard deviation was found to be well behaved (Officer, 1972). Later, Ang and Bekaert (2007) analyzed the ability of dividend yields to predict excess returns and cash flows. In their research, they found that these dividend yields were only effective short-term predictors, but lacked the ability to make long-term predictions.

To comprehend the impact of particular events on individual or multiple companies, MacKinlay (1997) employs an event study methodology. Event studies are a well-known tool for this purpose, not only in financial research but also in other fields. Event studies provide a systematic and quantitative approach to assessing the impact of specific events on financial markets. In order to do so, historic financial data is used to predict how an event impacted a company's stock market performance. Possible events range from corporate announcements such as mergers and acquisitions, earnings releases and changes in corporate governance to more general economic events such as interest rate changes, global pandemics or political changes.

Zeidan (2013) examined the impact of illegal behavior, specifically violations of laws and regulations, on the financial performance of U.S. banks. The authors used a sample size of 128 publicly traded banks subject to enforcement actions and an event study methodology, the author reports that these violations have a significant impact on the financial performance of banks. These market reactions even exceed the legal sanctions, which are also a real loss for the bank.

Bittlingmayer and Hazlett (2000) provide an example of the financial impact of antitrust enforcement initiatives against Microsoft. This article focuses on the impact of these enforcements on the computer industry. More specifically, it examines how the industry was affected by these charges and provides evidence against the belief that Microsoft's behavior is anticompetitive and that antitrust enforcement leads to overall efficiency gains. The study analyzed 54 announcements made between 1991 and 1997 to measure the stock price reactions of Microsoft and other computer industry firms, such as Western Digital, Apple and Seagate Technology.

Acquisti et al. (2006) have attempted to measure the impact of data breaches on a company's market value. According to their findings, data breaches have a negative and statistically significant impact on a company's market value on the day of disclosure. However, the impact was less than that observed in the literature for security breaches such as viruses, hacking or software vulnerabilities. The study also found that the effects of a data breach on a company's market value tends to be short-lived and that extreme market consequences are rarely observed.

Kleinow et al. (2014) analyzed the impact of regulatory announcements awarding a „systemic importance" seal to credit institutions on their stock market prices. They used an event study methodology to analyze stock price reactions to these announcements. The study found that market participants react to these regulatory announcements, but that stock returns are not exclusively positive. Furthermore, the market's response to the latest event appears to be decreasing, suggesting that the information's value is diminishing for those involved in the market.

Gao et al. (2011) conducted a study of stock and bond market responses to key events leading up to the Dodd-Frank Act. For the analysis of the stock market, 41 financial institutions were used, while for the analysis of the bond market, 31 institutions were used to estimate abnormal stock returns for 17 events. The study finds that there is an negative abnormal stock return and abnormal bond return. The positive impact on the bond market is due to the potential of the law to reduce the risk-taking of large banks.

Dangol (2008) analyzed the impact of announcements of unanticipated political events on the Nepalese stock market. An event study was used to examine how unanticipated political events such as the royal massacre, dissolution of parliament, Maoist activities and changes in government create political risk and uncertainty in the economy that affect stock prices. The results show that Nepalese investors reassess their stock prices when new political information is released. The study also found that the Nepalese stock market is not efficient, as the impact of new information is only observed there two or three days after the day of announcement. Finally, it was also found that bad news leads to stronger changes than good news. This finding is consistent with the views of Conrad et al. (2002), who report that stock prices are more sensitive to negative news than positive news.

Armour et al. (2017) examined the impact of regulatory sanctions on the market price of penalized firms in the United Kingdom. The study found that regulatory sanctions have a significant negative effect on the market price of penalized firms. The authors focused on the reputational losses and market reactions that occur after the announcements of regulatory sanctions on enforced firms. To measure the impact of these announcements on the market, an event study methodology was used. Furthermore, the authors conducted a cross-sectional analysis to investigate the impact of reputational sanctions. They found that reputational damage can be more severe than fines, especially when it affects customers or investors.

3. Hypotheses

The first research question of this master's thesis investigates whether there is a statistically significant difference in the distribution of cumulative abnormal returns for enforcement action announcements. The second research question investigates the extent to which variations in prediction can be observed using a cross-sectional analysis.

To answer these questions, an event study was conducted and two hypotheses were tested:

H1: The distribution of cumulative abnormal returns for announcements of enforcement actions is different from zero.

H2: Differences in prediction can be observed when using a cross-sectional analysis.

4. Materials & Methods

4.1. Initial Data

The initial datasets consist of enforcement actions issued by the FDIC, OCC and Fed. These datasets include information such as the date of issuance, the enforcement type, the URL of the legal notice, whether it is directed at an individual, the name of the individual and the name of the bank involved. They offer a complete overview of the regulatory actions enforced by these regulators and cover enforcement actions published from 1989 to 25 November 2023, the last date the data was accessed, comprising a total of 16,510 enforcement actions.

4.2. Data Processing

Standardization The initial stage involved standardizing the datasets by ensuring that all used features were brought into a common format. Furthermore, the types of enforcement were also standardized by renaming and only allowing certain types such as „Cease and Desist", „Written Agreement", „Section 19 Letter", „Prohibition", „Prompt Corrective Action", „Deposit Insurance Termination" and „Supervisory Agreement". Any types that did not fit into these categories were excluded, reducing the dataset to 15,968 actions. Figure 4.1 shows a sample of the standardized data.

date	type	url	isIndividual	bank	regulator
2009-11-25	Prohibition	/boarddocs/legaldevelopments/ordersother/secti...	1	wells fargo financial, inc., des moines, iowa,...	fed
2004-07-09	Written Agreement	/boarddocs/press/enforcement/2004/20040712/def...	0	first midwest bank, itasca, illinois	fed
2004-01-14	Civil Money Penalty	http://www.occ.gov/static/ots/enforcement/9358...	0	first state, fsb	occ
2011-06-06	Cease and Desist	https://orders.fdic.gov/sfc/servlet.shepherd/d...	0	freedom bank of america;	fdic
2008-08-15	Cease and Desist	https://orders.fdic.gov/sfc/servlet.shepherd/d...	0	greene county bank, the;	fdic

Figure 4.1.: Standardized Data

Name Matching The lack of an ISIN code in the regulatory data posed a considerable challenge in accurately identifying and comparing banks. To address this issue, a dataset of U.S. banks, including their names, ISIN codes and locations, was utilized. Initially, unnecessary words and characters such as „ of the", „ and", „„ inc", „ the", „ of", „ &", „;", „," and „." were removed from the names of both the banks in the regulatory data and the dataset. Irrelevant characters were removed prior to matching to ensure accuracy. The matching process involved a two-word matching approach, which proved to be more effective than using distance metrics such as Levenshtein distance or fuzzy matching techniques such as token, simple or partial ratios. Using this method, for instance, a bank such as „Bank of America" would transition to „Bank America" by eliminating the „ of" and subsequently be matched with an appropriate bank name from the other dataset. After matching, the dataset was reduced to 2651 entries due to the inability to match specific banks or the presence of foreign banks in the enforcement actions.

Refinitiv Eikon The distinct ISIN codes were extracted from the dataset to query the Refinitiv Eikon database. Four queries were conducted in total. The first query returned the stock returns, while the second included prices adjusted for holidays and weekends. The third investigated the S&P500 market return. Finally, the fourth query was a static query that retrieved a range of financial metrics, including market capitalization, return on assets and debt ratio. These statistics were gathered for the quarter prior to the relevant enforcement measure taken against the bank.

Abnormal Returns The following section utilized the event study approach to calculate the abnormal and cumulative abnormal returns of the enforced banks. As a result of the unavailability, incompleteness or inaccuracy of the stock data, the dataset was narrowed down to 746 instances.

Google Maps API To classify banks as situated either in the east or the west, the Google Maps API was utilized to obtain longitude and latitude data. Banks having a longitude below -100 were regarded as located in the east, whereas those with a longitude above -100 were classified as located in the west. Figure 4.2 shows the distribution of enforcement actions across the United States.

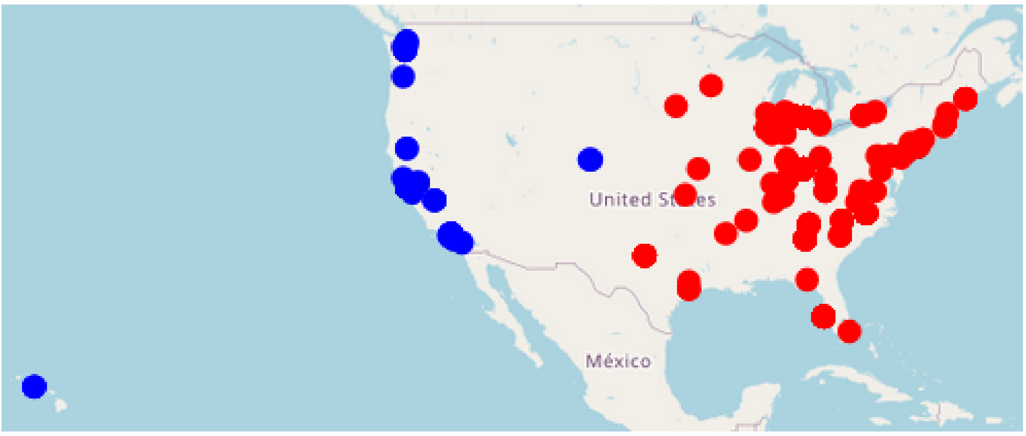

Figure 4.2.: Distribution of Enforced Banks

Legal Notice Text To tackle issues arising from broken and incomplete URLs leading to legal notices, measures to transform the relative or faulty URL were employed. First, I identified and removed URLs that were non-functional. Next, I reconstructed incomplete URLs to ensure they directed to the corresponding notice. Finally, I downloaded the first legal notice PDF that appeared upon accessing the link with the correct URL. The downloaded document was temporarily stored to enable PDF extraction before being subsequently deleted from the system. The text from the PDF was saved as a string feature for further processing. The dataset was reduced to 566 entries due to unreadable and absent PDFs.

NLP Preprocessing The extracted legal notice strings underwent further analysis and reduction in length involving the deletion of line breaks, tabs, links and unnecessary white space. The entire string was also converted to lower case for consistency. Repeated characters and punctuation were also removed to improve the accuracy and decrease the length of the analysis. Contraction words were expanded to their full forms and stop words were eliminated. In order to maintain correct spelling, a spell-checking procedure was applied, which also corrected errors found. Lastly, lemmatization was used to transform words in the text to their root form, enhancing the clarity and relevance of their meaning. This further minimizes the length of these sometimes relatively long legal texts and should help to minimize the cost of the large language models while maintaining the original meaning.

Same Day Enforcements To mitigate potential bias, enforcement actions taken against a single bank on the same day were aggregated and included as a feature. This approach aims to promote a more unbiased representation of enforcement activities.

OpenAI API The prepared legal notice strings were categorized using the OpenAI API, employing two legacy models: „gpt-3.5-turbo-0613" and „gpt-3.5-turbo-16k-0613". For strings with fewer than 4,000 tokens, the „gpt-3.5-turbo-0613" model was used, whereas the „gpt-3.5-turbo-16k-0613" model was used for longer ones, with up to 16,000 tokens. Running the data through the API for one iteration resulted in an approximate cost of 1.5$ per iteration.

To classify legal notices into four categories („low", „medium", „high" or „extreme"), the model's prompt was optimized using the OpenAI Playground. The model was first tested with explanations for each identified class and was later reduced to a single-word classification. A temperature of 0 was employed to regulate the level of randomness, resulting in more deterministic and repetitive outcomes. The final prompt used:

> *Imagine you are a U.S. lawyer specializing in enforcement actions related to the FED, OCC and FDIC against banks. Your task is to review a legal notice of an enforcement action and categorize it as „low", „medium", „high" or „extreme" risk to the bank's reputation. Please provide a one-word classification for the notice.*

The preprocessed legal notices were combined with the final prompt to query each legal text in the dataset, which allowed for a comprehensive analysis and categorization of the legal texts.

Generative AI Studio For comparison purposes, I employed the text-bison model from Google's Generative AI Studio. Note that this model did not have the capacity to process more than 8192 tokens. Hence, I removed the four enforcement actions that exceeded this limit. Adjusting the temperature to 0 managed the aspect of randomness, leading to results that were more deterministic and replicable. No charges were incurred as all models used were free at the moment of usage. To maintain consistency, I used the same prompt as for the OpenAI API. The classification process was kept similar to that of the OpenAI API.

Two tables summarizing the libraries and packages used in the study are attached in Appendix A. Furthermore, the language model's specific parameters are included in Appendix B.

4.3. Sensitivity Analysis

Extensive prompt engineering was carried out to develop the most appropriate prompts for both models, ensuring optimal classification accuracy. Initially, the models were asked to explain their classification, but this approach was finally changed to a single-word reply. A prompt template was created and evaluated for its effectiveness in achieving accurate classifications across different scenarios.

It was first attempted to categorize the legal texts into two groups, „severe“ and „not severe“, through the use of this prompt:

> *Imagine you are a U.S. lawyer specializing in enforcement actions related to the FED, OCC and FDIC against banks. Your task is to review a legal notice of an enforcement action and categorize it as „severe“ or „not severe“. Please provide a one-word classification for the notice.*

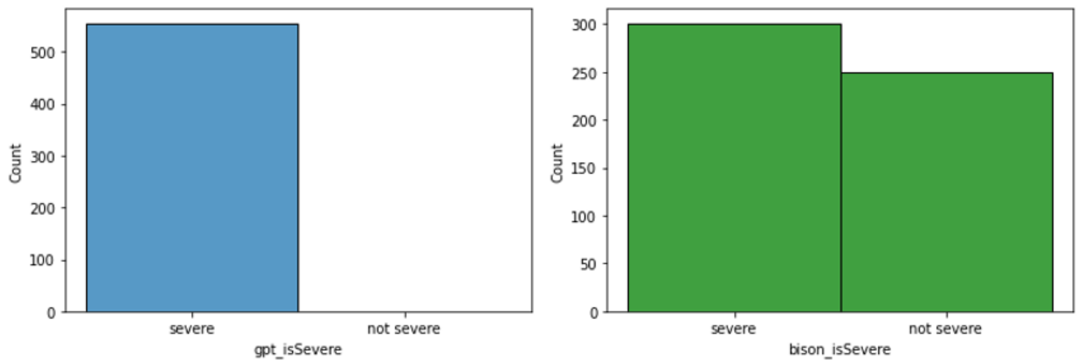

Figure 4.3.: Severe Classification by GPT and Bison

Figure 4.3 shows that the OpenAI model classifies all enforcements as severe, whereas the Google model has a more balanced distribution of classifications.

Using a different classification, the models were prompted to classify the notice into two different groups, namely „high“ and „low“ risk, using the following prompt:

> *Imagine you are a U.S. lawyer specializing in enforcement actions related to the FED, OCC and FDIC against banks. Your task is to review a legal notice of an enforcement action and categorize it as „low“ or „high“ risk to the bank's reputation. Please provide a one-word classification for the notice.*

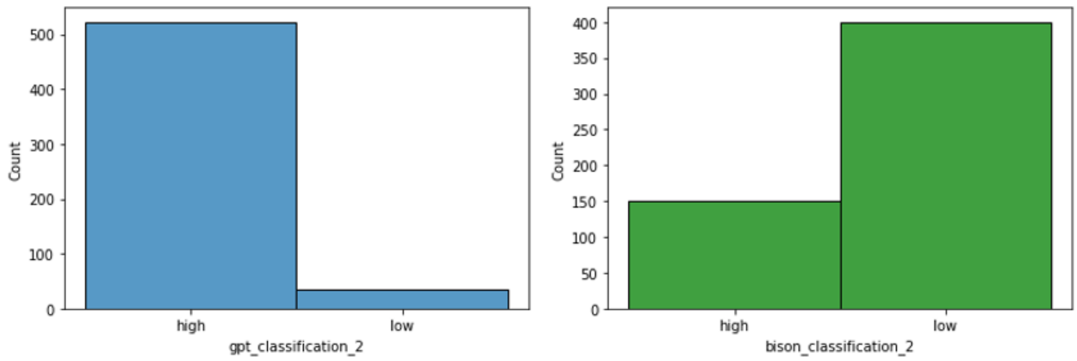

Figure 4.4.: Classification of 2 Classes by GPT and Bison

Figure 4.4 shows that the OpenAI model classified most notices as high risk, while the Google model was more generous, classifying the majority as low risk.

The final model was prompted to classify the texts into four classes, namely „extreme", „high", „medium" and „low", using the provided prompt:

> *Imagine you are a U.S. lawyer specializing in enforcement actions related to the FED, OCC and FDIC against banks. Your task is to review a legal notice of an enforcement action and categorize it as „low", „medium", „high" or „extreme" risk to the bank's reputation. Please provide a one-word classification for the notice.*

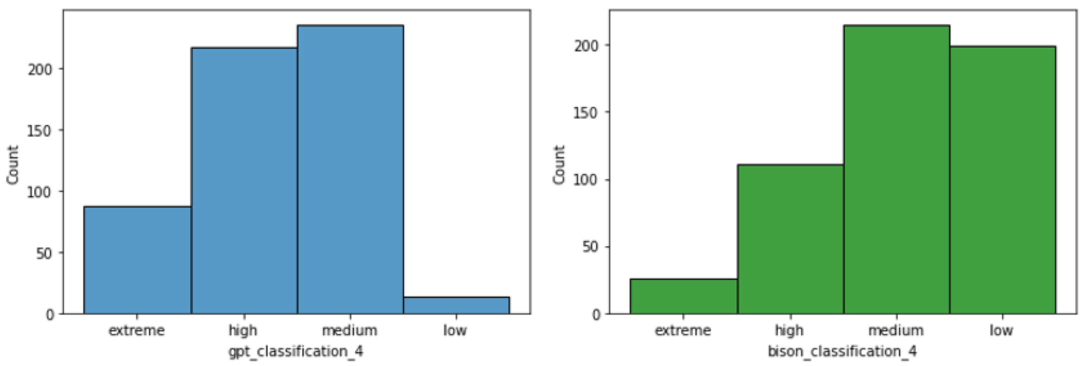

Figure 4.5.: Classification of 4 Classes by GPT and Bison

Figure 4.5 once again highlights the difference in classification between the OpenAI and Google models when it comes to legal notices. While GPT model's least common classification was „low", Bison model's least common classification was „extreme" with a significant portion of notices falling into the „low" risk category. These findings highlight the differences in the two models when classifying legal texts.

4.4. Event Study Approach

An event study methodology has been used to measure abnormal stock returns in enforcement actions against listed U.S. banks. Given rationality in the market, the assumption of this methodology is that financial markets react to enforcement actions, so that the stock price is a good parameter to measure the impact of these actions. Following the event study methodology highlighted by MacKinlay (1997), the first step is to estimate normal returns.

4.4.1. Estimating Normal Returns

There are two main statistical approaches to estimating the normal return of a given security: the constant mean return model and the market model. The common assumption of these models is that they follow a jointly multivariate normal distribution and remain independent and identically distributed over time. This underlying assumption is usually not problematic as it is consistent with empirical observations.

Constant Mean Return Model As its name suggests, the constant mean return model assumes that the average return of a specific security remains consistent over time. Brown and Warner (1980) mentioned that although the constant mean return model is the simpler model, it often produces results that do not differ from more complex models.

$$R_{it} = \mu_i + \zeta_{it} \tag{4.1}$$

where:

R_{it} = the return of stock i on day t
μ_i = the mean return for stock i
ζ_{it} = the disturbance term of stock i on day t

Market Model The market model offers potential improvements over the constant mean return model by eliminating some of the variability associated with fluctuations in market returns. It predicts the performance of a security relative to the performance of a market portfolio by assuming a stable linear relationship between the two, which follows from the assumption of joint normality of asset returns. In addition, the market model is the most commonly used method with good predictive power (Brenner, 1979).

$$R_{it} = \alpha_i + \beta_i R_{mt} + \varepsilon_{it} \tag{4.2}$$

where:

R_{it} = the return of stock i on day t
R_{mt} = the return of market portfolio m on day t
α_i *and* β_i = market model parameters
ε_{it} = the prediction error of stock i on day t

With a normal return model chosen, the next step is to determine the estimation window. The most common approach is to use the period preceding the event window as the estimation window. For each enforcement announcement, the model is estimated using 250 daily returns ending before the event window and is defined by the period $T_0 + 1$ to T_1 (see Figure 4.6). The estimation and event windows do not overlap to avoid the event potentially influencing the estimate of the normal return.

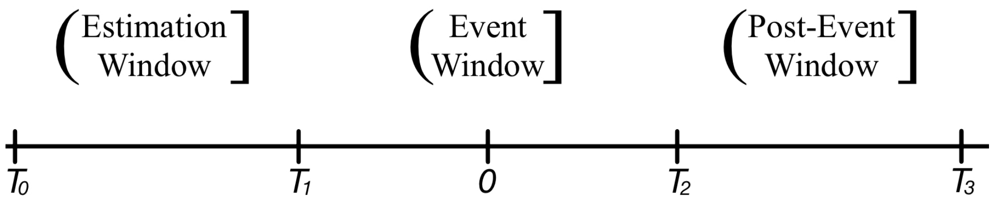

Figure 4.6.: Event Periods Timeline

4.4.2. Measuring Abnormal Returns

Using the normal returns predicted by the market or constant mean return model, the abnormal return (AR) can be calculated. The *AR* for stock i during event time τ is calculated by subtracting the actual share returns from its expected stock returns. This represents the realized return that was not explained by market movements and is designed to capture the impact of the event. Consequently, it serves as the representation of prediction errors for stock i during event time τ:

$$AR_{i\tau} = R_{i\tau} - (\hat{\alpha}_i + \hat{\beta}_i R_{m\tau}) \tag{4.3}$$

Cumulative Abnormal Returns (CAR) are calculated to aggregate the abnormal returns over an event window and draw general conclusions. Typically, the event window is configured to extend beyond a single day. This approach ensures that the impact of announcements made after market close is captured. At the same time, the inclusion of the day before the event captures the impact of information leaked prior to the announcement. The calculation of CAR involves aggregating the abnormal returns for a single stock i over the event window $T_1 + 1$ to T_2 (see Figure 4.6):

$$CAR_i(\tau_1, \tau_2) = \sum_{\tau=\tau_1}^{\tau_2} AR_{i\tau} \tag{4.4}$$

where:

$\tau_1 =$ the beginning time point of the analysis
$\tau_2 =$ the ending time point of the analysis

Average abnormal returns (AAR) serve the purpose of aggregating abnormal returns across multiple events. It assumes the absence of clustering, indicating non-overlapping event windows. As a result, it assumes the independence of abnormal returns and cumulative abnormal returns. By employing formula 4.3 and considering a total of N events, aggregated abnormal returns can be calculated:

$$\overline{AR}_\tau = \frac{1}{N} \sum_{i=1}^{N} AR_{i\tau} \tag{4.5}$$

$$var(\overline{AR}_\tau) = \frac{1}{N^2} \sum_{i=1}^{N} \sigma^2_{\epsilon_i} \tag{4.6}$$

Finally, in order to aggregate abnormal return across both securities and over time, the concept of cumulative average abnormal returns (CAAR) comes into play. This method employs the same approach outlined in formula 4.4 utilizing the AAR:

$$\overline{CAR}_i(\tau_1, \tau_2) = \sum_{\tau=\tau_1}^{\tau_2} \overline{AR}_\tau \tag{4.7}$$

$$var(\overline{CAR}_i(\tau_1, \tau_2)) = \sum_{\tau=\tau_1}^{\tau_2} var(\overline{AR}_\tau) \tag{4.8}$$

4.5. Hypothesis Testing

Due to the assumption of non-clustering, the covariance term is set to 0. As a result, considering the null hypothesis H_0, which suggests that the events have no impact on the return of a security, the distribution of CAAR can be drawn using:

$$\overline{CAR}(\tau_1, \tau_2) \sim \mathcal{N}[0, var(\overline{CAR}(\tau_1, \tau_2))] \tag{4.9}$$

Utilizing the distribution outlined in formula 4.9 alongside the CAAR enables the testing of the null hypothesis assuming that abnormal returns are equal to zero. Since the variance $\sigma^2_{\epsilon_i}$ is usually unknown, an estimator is used to calculate the variance of abnormal returns. Using these measures the null hypothesis can be evaluated by:

$$\theta_1 = \frac{\overline{CAR}(\tau_1, \tau_2)}{var(\overline{CAR}(\tau_1, \tau_2))^{\frac{1}{2}}} \sim \mathcal{N}(0, 1) \tag{4.10}$$

4.6. Multiple Linear Regression

Multiple linear regression is a statistical method that allows predictions using multiple predictor variables. It extends the previously used simple linear regression model by allowing for the inclusion of more than one predictor variable. The assumption that there is a linear relationship between the response variable and each predictor variable is maintained (Tranmer and Elliot, 2008). Furthermore, it is assumed that the variables have constant variance, are independent and are normally distributed (Eberly, 2007).

The following equation is used to predict and analyze the impact of the predictor variables on the response variable for i = 1, 2,. . . , n instances:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_{p-1} x_{ip-1} + \varepsilon_i \tag{4.11}$$

where:

β_0 = mean Y when predictors are 0 (intercept)
β_k = rate of change (slope)

The following table presents the features used in the linear regression analysis:

Feature	Description
bison_classifier	Classification by Generative AI Studio
gpt_classifier	Classification by OpenAI
count_day_eas	# Enforcements for the same Bank on the same Day
isEast	East/West Location Indicator
isIndividual	Individual/Bank Indicator
mkt_cap	Total Stock Market Value
roa	Metric for Company's Profitability
debt_ratio	Metric for Capital Structure

Table 4.1.: Feature Table

The model is evaluated using several techniques. Firstly, the coefficients of the predictor variables are examined. These coefficients indicate the change in the response variable for a one-unit change in the predictor variable. Then the R-squared value is examined, which measures the proportion of variance in the response variable that is explained by the model. A value of 1 indicates a perfect linear dependence, whereas a value of 0 indicates that the model could not explain the variance in the response variable. Finally, the residuals are analyzed for constant variance (heteroscedasticity), which is assumed by a linear regression.

4.7. Random Forest Regression

Random forests is a machine learning technique that can be used to predict categorical variables and perform non-linear regression tasks. By combining multiple tree predictors that depend on a random vector, random forests can provide more accurate predictions than linear regression models. The following method is based on Breiman (2001).

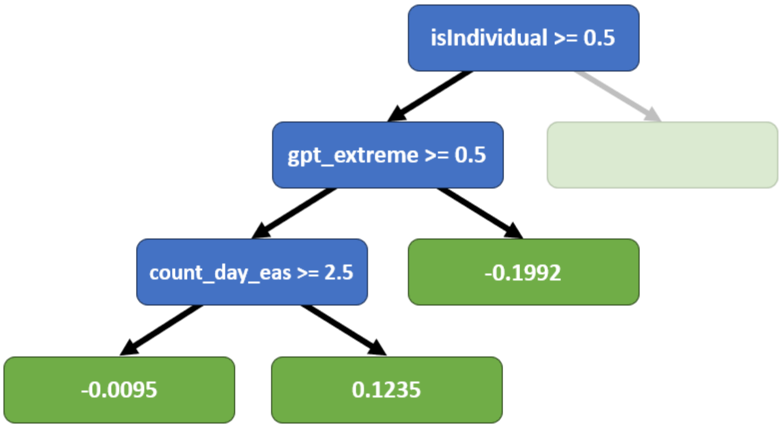

Figure 4.7.: Example Decision Tree

The concept of random forests for regression is similar to the traditional version used for classification. In this approach, bagging (Bootstrap Aggregating) is employed to train multiple trees on various bootstrap samples of the training data. Here, samples are randomly selected with replacement from the initial training set. Trees (see Figure 4.7) are then constructed using a random vector Θ and the predictor $h(x, \Theta)$ uses numerical values instead of classifications to make predictions. In contrast to the classical random forests approach, where classification is determined by majority voting, in regression tasks the predictor is obtained by averaging the predictions of all trees. This allows for more accurate predictions and is less prone to overfitting due to the law of large numbers and the averaging of multiple trees helps to reduce the variance of the predictions.

5. Evaluation

This section of the master's thesis presents the empirical results obtained from the conducted event study, as well as a comparison between the two large language models used from OpenAI and Google.

5.1. Results

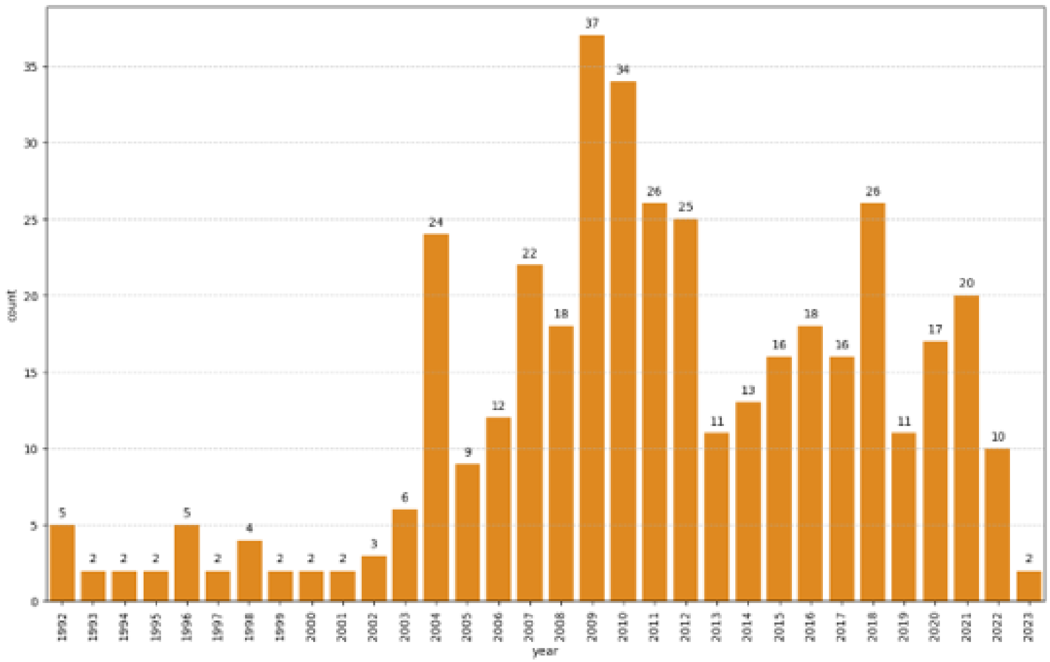

Figure 5.1.: Number of Enforcement Actions per Year

Figure 5.1 shows for the final dataset the distribution of enforcement actions issued by the OCC, FDIC or Fed from 1992 to 2023. Each bar represents a specific time period, with the numbers above indicating the total count of enforcement actions issued. Looking at the number of actions taken over the years, it is evident that the number of enforced bank regulations has increased. It is also worth noting that there was a significant increase in enforcements during the financial crisis in 2008.

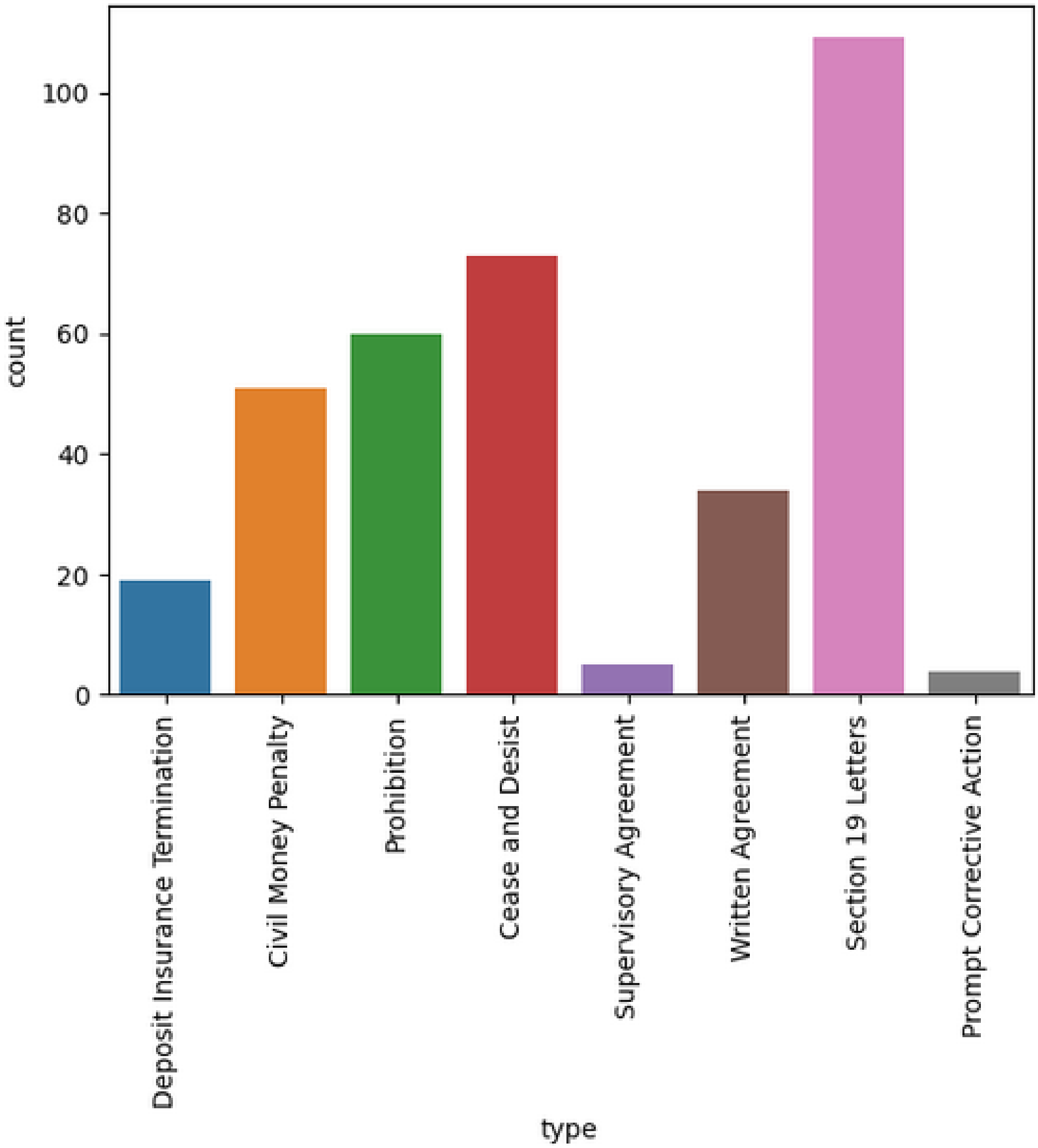

Figure 5.2.: Number of Enforcements per Type

Figure 5.2 illustrates for the final dataset the distribution of enforcement types issued by the OCC, FDIC or Fed from 1992 to 2023. The dataset shows that „Section 19 Letters" is the most frequently used type of enforcement, while „Supervisory Agreement" and „Prompt Corrective Action" are rarely used. In contrast, „Cease and Desist", „Prohibition", „Civil Money Penalty" and „Written Agreement" are regularly employed as enforcement types.

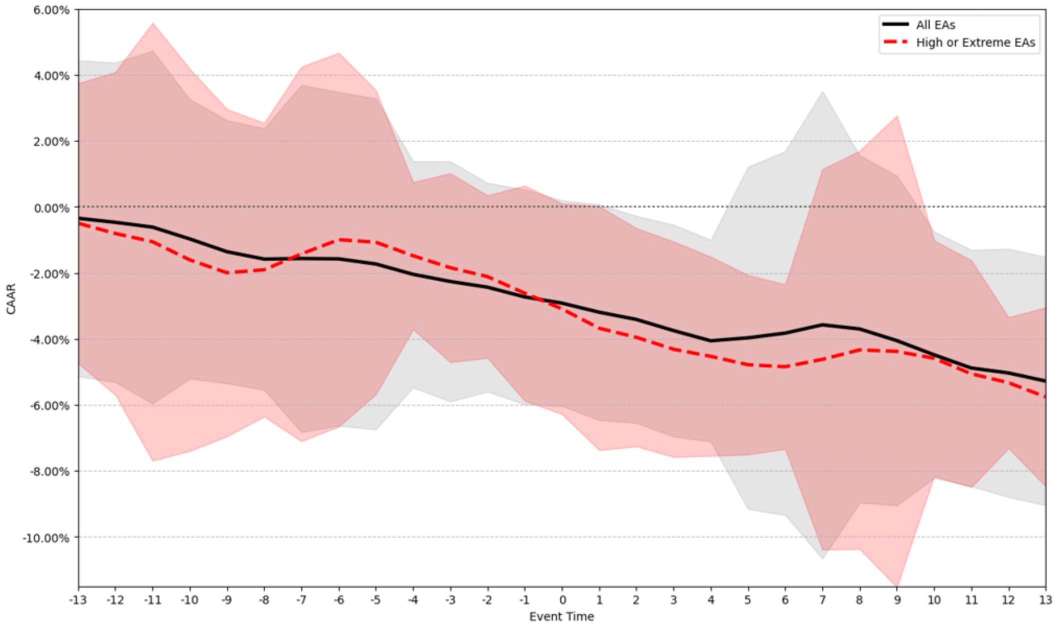

Figure 5.3.: Cumulative Average Abnormal Returns over Time

Figure 5.3 displays the cumulative average abnormal returns for enforced banks over a 26-day period. The returns were calculated by averaging the returns from 13 days before the announcement to 13 days after using a 3-day rolling average. The black solid line in the figure represents the average abnormal return for all enforcement actions, while the red dotted line represents enforcement actions classified as „high" or „extreme" by the Google API. The variance around the calculated means is indicated by the black and red shadow.

The average abnormal returns for all enforcements, as well those returns averaged for the enforcements classified as „high" or „extreme", were already negative prior to the announcement. However, the returns began to consistently decline approximately 6 days before the announcements and the more extreme enforcements experienced even sharper drops after the announcements. Although there was a slight rebound about 6 days after announcement, the average returns continued the negative trend. The calculated returns appear to be less variable during this period.

To test the hypothesis that the distribution of cumulative abnormal returns for announcements of enforcement actions differs from zero, a one-sided t-test was conducted to determine whether the distribution of cumulative abnormal returns is different from 0. The market model was employed for this analysis. Due to the unavailability of data for windows, the number of events varies across different windows. Throughout this thesis, the following significance codes are used: „***“ for p-values below 0.001, „**“ for p-values below 0.05, „*“ for p-values below 0.1 and „.“ for p-values below 0.1.

Sample	N	Mean (%)	t-score
CAR(-1,1)	404	-0.69	-2.552*
CAR(-3,3)	361	-1.05	-2.32*
CAR(-5,5)	314	-1.07	-1.981*
CAR(-1,3)	364	-0.54	-1.216.
CAR(-3,1)	394	-1.26	-3.463***

Table 5.1.: Test Scores

Table 5.1 shows that all event windows displayed negative mean cumulative abnormal returns. The event window from 3 days before to 1 day after the event showed the highest significance with a p-value below 0.001, while the window from 1 day before to 3 days after the event had the lowest significance. The findings support the hypothesis that the distribution of cumulative abnormal returns for announcements of enforcement actions is different from zero.

5.1.1. Cross-Sectional Analysis

To test the second hypothesis, multiple models were implemented to observe differences in prediction when using a cross-sectional analysis. Furthermore, additional business performance indicators were adopted as features from Pugachev and Schertler (2021). First, the natural logarithm market capitalization (mkt_cap) is a metric used to determine a company's overall value in the stock market and a snapshot of the size of a company. The calculation is made by multiplying the current price of the share by the total number of shares in circulation. Second, the return on assets (roa) is a financial metric used to assess a company's profitability. It is calculated by dividing the net income generated by the company's assets, which helps to determine how effectively the company is utilizing its resources to generate profits. Lastly, the debt ratio (debt_ratio) is a financial metric that evaluates a company's capital structure. It measures the percentage of a company's total debt to its total assets, providing insight into the level of financial leverage and risk.

Before doing so, the cumulative abnormal returns were winsorized at 95% to avoid having highly influential points in the data. Furthermore, I excluded legal notices that were classified as „low“ by the OpenAI or as „extreme“ by the Google model when using the classifier feature. This was done due to the fact that there were only a few legal notices classified as such. Due to lack of space, I further excluded the window that captures effects from five days before to five days after the announcement.

Multiple Linear Regression – Market Model

	CAR (-1,1)	CAR (-1,1)	CAR (-3,3)	CAR (-3,3)	CAR (-1,3)	CAR (-1,3)	CAR (-3,1)	CAR (-3,1)
(Intercept)	-0.001 (0.011)	-0.004 (0.011)	-0.032. (0.018)	-0.044* (0.019)	-0.025. (0.014)	-0.026. (0.016)	-0.013 (0.014)	-0.021 (0.015)
bison_low		0.007 (0.005)		0.02** (0.008)		0.015* (0.006)		0.012* (0.006)
bison_medium		0.008. (0.004)		0.015* (0.007)		0.008 (0.006)		0.011* (0.006)
gpt_medium	0.004 (0.004)		0.011 (0.008)		0.012. (0.006)		0.005 (0.006)	
gpt_high	0.003 (0.005)		0.004 (0.008)		0.001 (0.006)		0.006 (0.006)	
count_day_eas	0.001 (0.003)	0.002 (0.003)	0.012* (0.006)	0.014* (0.006)	0.008. (0.005)	0.009. (0.005)	0.003 (0.004)	0.005 (0.004)
isEast	-0.004 (0.004)	-0.004 (0.004)	0.009 (0.007)	0.009 (0.007)	0.006 (0.005)	0.005 (0.005)	0.001 (0.005)	0.001 (0.005)
isIndividual	0 (0.004)	0.001 (0.003)	0.005 (0.006)	0.007 (0.006)	0.004 (0.005)	0.004 (0.005)	0 (0.005)	0.003 (0.004)
mkt_cap	0 (0.001)	-0.001 (0.001)	0 (0.001)	0 (0.001)	0 (0.001)	-0.001 (0.001)	0 (0.001)	0 (0.001)
roa	0.007* (0.003)	0.007* (0.003)	-0.003 (0.006)	-0.004 (0.006)	-0.003 (0.005)	-0.003 (0.005)	0.007 (0.004)	0.006 (0.004)
debt_ratio	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0. (0)	0* (0)
N	393	387	351	345	353	348	384	378
F-statistic	1.448	1.930.	1.303	1.894.	1.298	1.378	1.759.	2.271*

Table 5.2.: Cross-Sectional Regressions for Market Model
showing coefficients with standard errors in brackets

Table 5.2 shows the results of a multiple linear regression analysis using the market model as a reference. Columns 1 and 2 demonstrate the models for the (-1,1) window, comparing the addition of OpenAI features with Google features. The same approach was used for the (-3,3), (-1,3) and (-3,1) windows. The model displays the estimates, their significance codes and standard errors at each cell.

It is expected that more severe actions will lead to more negative stock returns. The intercepts in the models indicate the highest level of severity with GPT classified as „extreme“ and bison classified as „high“. The analysis indicates that in these cases, cumulative returns show more notable changes and lower classifications tend to result in less negative returns. However, it is important to note that although some models are statistically significant, their low R-squared values indicate that they only explain a small portion of the variability in the dependent variable. Based on the residuals, the model satisfies the assumption of homoscedasticity.

Multiple Linear Regression – Constant Mean Model

	CAR (-1,1)	CAR (-1,1)	CAR (-3,3)	CAR (-3,3)	CAR (-1,3)	CAR (-1,3)	CAR (-3,1)	CAR (-3,1)
(Intercept)	-0.009 (0.012)	-0.009 (0.013)	-0.035. (0.021)	-0.062** (0.023)	-0.028. (0.016)	-0.042* (0.017)	-0.03. (0.016)	-0.036* (0.017)
bison_low		0.005 (0.005)		0.03*** (0.009)		0.018** (0.006)		0.014* (0.007)
bison_medium		0.009. (0.005)		0.026** (0.008)		0.013* (0.006)		0.018** (0.007)
gpt_medium	0.004 (0.005)		0.009 (0.009)		0.006 (0.007)		0.009 (0.007)	
gpt_high	0.002 (0.005)		0.001 (0.009)		-0.002 (0.007)		0.006 (0.007)	
count_day_eas	0.005 (0.003)	0.005. (0.003)	0.019** (0.007)	0.022** (0.007)	0.016** (0.005)	0.019*** (0.005)	0.01* (0.004)	0.011* (0.004)
isEast	-0.002 (0.005)	-0.003 (0.004)	0.008 (0.008)	0.009 (0.008)	0.007 (0.006)	0.007 (0.006)	0 (0.006)	-0.001 (0.006)
isIndividual	-0.003 (0.004)	-0.004 (0.004)	0.004 (0.007)	0.006 (0.007)	0.003 (0.005)	0.004 (0.005)	-0.001 (0.006)	0 (0.005)
mkt_cap	0 (0.001)	0 (0.001)	0 (0.001)	0.001 (0.001)	0 (0.001)	0 (0.001)	0.001 (0.001)	0.001 (0.001)
roa	0.005 (0.004)	0.006 (0.004)	-0.007 (0.007)	-0.008 (0.007)	-0.011* (0.005)	-0.012* (0.005)	0 (0.005)	0 (0.005)
debt_ratio	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0. (0)	0. (0)
N	393	387	351	345	353	348	384	378
F-statistic	1.166	1.670.	1.378	2.739**	2.476.	3.463**	1.748.	2.267*

Table 5.3.: Cross-Sectional Regressions for Constant Mean Model
showing coefficients with standard errors in brackets

Table 5.3 presents results comparable to those in Table 5.2, using the constant mean model to calculate abnormal returns.

5.1.2. Random Forest Regressor

To perform the forest regression, a random sample of 20% of the data was reserved as test data. The remaining data (training data) was used to train a random forest model with 1000 trees. The model was then tested using the 20% test data to calculate errors. Although the model has a relatively low mean squared error of 0.0005, it produces negative R-squared values, which could be due to a number of reasons. First, if the random forest model is not suitable for the data or task provided, it may perform poorly overall, indicating that a simpler model might be more appropriate. Furthermore, the limited size of the dataset may worsen the problem by increasing the likelihood of statistical variation. However, this is not surprising as the previous models had low R-squared values, indicating that they might have been unable to extract much insight from the selected features.

6. Conclusion

This thesis examined the information content of enforcement actions against U.S. based banks. A review of the literature shows that there is some evidence that enforcement actions affect the share prices of enforced banks, as well as other aspects of their operations. As there was only a small amount of literature focusing on event studies of enforcement actions, the aim of this paper was to fill this gap, while also integrating nowadays relevant large language models, like the ones from OpenAI or Google.

An event study was conducted to calculate the cumulative abnormal returns for the different enforcements and event windows. These calculated returns were positively tested for significance, with the most significant event window, which also had the most negative average return, starting 3 days before and ending 1 day after the announcement. This may be due to the fact that these enforcement actions often do not come as a surprise, but some information is often known before the actual announcement. The extent to which investors have prior knowledge remains unknown and might be subject to further research. To examine these returns further, several models were employed to try to explain the variance in the predictions. Based on the results of these models, it can be said that they did not contribute much to the explanation of the variance.

In summary, it can therefore be stated that the enforcement actions against U.S. banks by the OCC, FDIC and Fed did have a significant impact on the corresponding stock market price, but the thesis failed to gain any further insights. Regarding the language models used, they could be helpful in classifying legal texts, but this will need additional research to be confirmed. Finally, during the course of my research I was confronted with a number of issues that could further limit the findings. Firstly, the bank name matching could have led to false matches. Additionally, the URLs leading to legal notices occasionally included multiple legal texts. For simplification, only the first legal text was extracted, which may have resulted in the omission of important information.

Bibliography

Acquisti, A., Friedman, A., & Telang, R. (2006). Is there a cost to privacy breaches? an event study. *ICIS 2006 proceedings*, 94 (cit. on p. 8).

Ang, A., & Bekaert, G. (2007). Stock return predictability: Is it there? *The Review of Financial Studies*, *20*(3), 651–707 (cit. on p. 7).

Armour, J., Mayer, C., & Polo, A. (2017). Regulatory sanctions and reputational damage in financial markets. *Journal of Financial and Quantitative Analysis*, *52*(4), 1429–1448 (cit. on p. 9).

Bittlingmayer, G., & Hazlett, T. W. (2000). Dos kapital: Has antitrust action against microsoft created value in the computer industry? *Journal of Financial Economics*, *55*(3), 329–359 (cit. on p. 8).

Breiman, L. (2001). Random forests. *Machine learning*, *45*, 5–32 (cit. on p. 22).

Brenner, M. (1979). The sensitivity of the efficient market hypothesis to alternative specifications of the market model. *The Journal of Finance*, *34*(4), 915–929 (cit. on p. 17).

Brous, P. A., & Leggett, K. (1996). Wealth effects of enforcement actions against financially distressed banks. *Journal of Financial Research*, *19*(4), 561–577 (cit. on p. 6).

Brown, S. J., & Warner, J. B. (1980). Measuring security price performance. *Journal of financial economics*, *8*(3), 205–258 (cit. on p. 17).

Conrad, J., Cornell, B., & Landsman, W. R. (2002). When is bad news really bad news? *The Journal of Finance*, *57*(6), 2507–2532 (cit. on p. 9).

Dangol, J. (2008). Unanticipated political events and stock returns: An event study. *NRB Economic Review*, 86–110 (cit. on p. 9).

Danisewicz, P., McGowan, D., Onali, E., & Schaeck, K. (2018). The real effects of banking supervision: Evidence from enforcement actions. *Journal of Financial Intermediation*, *35*, 86–101 (cit. on p. 7).

Delis, M. D., Staikouras, P. K., & Tsoumas, C. (2017). Formal enforcement actions and bank behavior. *Management Science*, *63*(4), 959–987 (cit. on p. 6).

Eberly, L. E. (2007). Multiple linear regression. *Topics in Biostatistics*, 165–187 (cit. on p. 21).

FDIC. (2023a). *About the fdic*. https://www.fdic.gov/about/. (Cit. on p. 4)

FDIC. (2023b). *Fdic enforcement decisions and orders*. https://orders.fdic.gov/s/types-of-action. (Cit. on p. 6)

Fed. (2023a). *About the fed*. https://www.federalreserve.gov/aboutthefed.htm. (Cit. on p. 4)

Fed. (2023b). *Enforcement actions about*. https://www.federalreserve.gov/supervisionreg/enforcement-actions-about.htm. (Cit. on p. 5)

Gao, Y., Liao, S., & Wang, X. (2011). The economic impact of the dodd frank act of 2010: Evidence from market reactions to events surrounding the passage of the act. *Social Sciences Research Network* (cit. on p. 8).

Jordan, J. S., Peek, J., & Rosengren, E. S. (1999). *The impact of greater bank disclosure amidst a banking crisis* (Vol. 2). Federal Reserve Bank of Boston. (Cit. on p. 7).

Kleinow, J., Nell, T., Rogler, S., & Horsch, A. (2014). The value of being systemically important: Event study on regulatory announcements for banks. *Applied Financial Economics*, *24*(24), 1585–1604 (cit. on p. 8).

MacKinlay, A. C. (1997). Event studies in economics and finance. *Journal of economic literature*, *35*(1), 13–39 (cit. on pp. 7, 17).

OCC. (2023a). *About the occ*. https://www.occ.treas.gov/about/index-about.html. (Cit. on p. 3)

OCC. (2023b). *Enforcement action types*. https://www.occ.treas.gov/topics/laws-and-regulations/enforcement-actions/enforcement-action-types/index-enforcement-action-types.html. (Cit. on pp. 5, 6)

Officer, R. R. (1972). The distribution of stock returns. *Journal of the american statistical association*, 807–812 (cit. on p. 7).

Peek, J., & Rosengren, E. (1995). Bank regulation and the credit crunch. *Journal of Banking & Finance*, *19*(3-4), 679–692 (cit. on p. 6).

Pugachev, L., & Schertler, A. (2021). Neglecting peter to fix paul: How shared directors transmit bank shocks to nonfinancial firms. *Journal of Financial and Quantitative Analysis*, *56*(6), 2170–2207 (cit. on p. 26).

Tranmer, M., & Elliot, M. (2008). Multiple linear regression. *The Cathie Marsh Centre for Census and Survey Research (CCSR)*, *5*(5), 1–5 (cit. on p. 21).

Zeidan, M. J. (2013). Effects of illegal behavior on the financial performance of us banking institutions. *Journal of business ethics*, *112*, 313–324 (cit. on p. 7).

Appendix

Appendix A.

Used Libraries

Library	Version
autocorrect	2.6.1
backoff	2.2.1
bs4	0.0.1
datetime	5.4
folium	0.15.0
googlemaps	4.10.0
matplotlib	3.7.2
nltk	3.7
numpy	1.23.2
openai	0.27.2
pandas	2.0.3
PyPDF2	3.0.1
requests	2.31.0
scipy	1.11.2
seaborn	0.12.2
sklearn	1.3.1
tiktoken	0.5.1
unidecode	1.3.6
vertexai	1.39.0

Figure A.1.: Python Libraries Table

Library	Version
dplyr	1.0.9
lubridate	1.9.0
readr	2.1.2
reshape2	1.4.4
rlang	1.0.4
rpart	4.1.16
tibble	3.1.8

Figure A.2.: R Libraries Table

Appendix B.

LLM Settings

openai.ChatCompletion.create()
model = gpt-3.5-turbo-0613/gpt-3.5-turbo-16k-0613 temperature = 0 n = 1 top_p = 1 messages = [{"role": "system", "content": final_prompt}, {"role": "user", "content": legal_notice}]

Table B.1.: OpenAI Parameters

TextGenerationModel.from_pretrained("text-bison").predict()
temperature = 0 candidate_count = 1 max_output_tokens = 100 top_p = 1 prompt = final_prompt + row[legal_notice]

Table B.2.: Generative AI Studio Parameters