Nijmegen School of Management
Department of Economics and Business Economics
Master's Thesis in Economics (MAN-MTHEC)

The Moderating Role of Corporate Social Responsibility in Market Reactions to Financial Misconduct

By Nick Veenhuizen (S1012716)

Nijmegen, 30 June 2025

Program: Accounting & Control
Supervisor: dr. E.K. Matthaei

Radboud Universiteit

No Generative AI tools were used during the development and writing of this thesis. By submitting this thesis I declare that I am fully responsible for the accuracy and completeness of its content.

ABSTRACT

This thesis investigates whether Corporate Social Responsibility (CSR) moderates investor responses to announcements of financial misconduct, including tax violations, anti-money laundering deficiencies, and accounting fraud. Using an event study methodology, it examines short-term cumulative abnormal returns (CARs) for a sample of 80 violation cases and 157 matched control firms. Firms were matched on financial and industry characteristics using a propensity score matching algorithm implemented with the MatchIt package in R. CSR performance was measured using the Refinitiv ESG Combined Score, and firms were categorized into tiers, focusing on the top 25% (HighCSR) for regression analyses. The results show no statistically significant market reactions to the misconduct events, nor a consistent moderating effect of CSR. A power analysis confirms low statistical power to detect small effects, particularly for interaction terms. These findings show the empirical limitations in capturing the reputational role of CSR and suggest that future studies require larger samples and refined CSR measures.

Table of Contents

1 Introduction

Corporate tax behaviour has become a topic of growing concern for investors, regulators, and society at large. Over the past two decades, public awareness and scrutiny of tax avoidance have steadily increased, driven by high-profile scandals such as the Panama Papers and stronger calls for corporate accountability (Graham, Hanlon & Shevlin, 2014). Stakeholders now expect firms not only to comply with tax laws but also to uphold ethical standards and contribute fairly to public finances (Hanlon & Heitzman, 2010). This shift in expectations has made tax behaviour an important factor in a company's reputation and long-term stakeholder relationships.

While firms often defend tax planning as a legal way to maximize shareholder value, research shows that the line between acceptable strategies and aggressive avoidance can be blurry and controversial. Sikka (2010) argues that companies frequently present themselves as socially responsible while exploiting tax loopholes at the same time. This contradiction undermines the credibility of their wider CSR efforts and can decrease public trust. Similarly, Austin and Wilson (2017) find that firms with well-known consumer brands face higher reputational risks when they engage in aggressive tax avoidance. Investors and customers are more likely to view such behaviour as hypocritical, especially when a company's public image is built on ethical claims.

As a result, tax behaviour is no longer just a technical financial decision. It has become a reputational risk factor that can affect investor confidence and market value, especially when misconduct is revealed. However, in practice, tax violations alone occur relatively infrequently and are often part of broader issues related to financial reporting and compliance failures. For this reason, this study includes not only tax violations but also cases of anti-money-laundering deficiencies and accounting fraud or reporting misconduct. These types of violations share common features: they weaken a firm's financial transparency, signal weak governance, and pose similar reputational risks (Davidson, Worrell & Lee, 1994; Ichev, 2023). Combining them into one category ensures a sufficient number of observations for empirical testing while remaining consistent with the theoretical focus on financial misconduct that affects stakeholder trust.

When a company's financial misconduct becomes public knowledge, this disclosure can significantly damage its reputation and reduce investor trust. Such announcements often signal deeper problems with corporate governance and ethics. Kim, Li, and Zhang (2011) show that firms that use aggressive tax strategies tend to have less transparent financial reporting. This lack of transparency allows managers to hide negative information for longer. When the hidden issues eventually surface, the sudden revelation can cause sharp drops in share prices, known as stock price crashes.

Recent research also highlights the important role of media coverage in shaping investor reactions to corporate misconduct. Ichev (2023) shows that the tone and framing of news reports can either soften or amplify the market's response. When articles use strong negative language and clearly blame the company instead of external factors or individuals, investors tend to react more strongly. This effect is especially visible in cases of financial misconduct such as tax fraud or accounting scandals. Firms involved in such controversies often face sharper stock price drops when media coverage highlights corporate blame and ethical failings.

Earlier work by Davidson, Worrell, and Lee (1994) provides further evidence that the market punishes firms for corporate crimes, including tax evasion, bribery, and financial fraud. By analysing over 500 cases of corporate wrongdoing, they show that illegal actions not only lead to direct financial penalties but also damage a firm's reputation in the eyes of investors. Companies with a history of misconduct face even stronger negative reactions, as stakeholders may see new violations as part of a larger pattern of poor governance.

In this context, Corporate Social Responsibility (CSR) is often viewed as a way to reduce reputational risks. A strong CSR profile can signal to stakeholders that a firm acts responsibly and cares about social and environmental issues beyond profit (Godfrey, Merrill, & Hansen, 2009). This reputation builds "moral capital" that can protect a company when problems arise. Minor and Morgan (2011) support this view by showing that firms with strong CSR records face smaller negative market reactions during product recalls. However, not all research agrees that CSR always shields firms from harm. Bartov et al. (2019) find that when wrongdoing is seen as intentional or fraudulent, a high CSR score can actually increase investor disappointment. In such cases, stakeholders may feel misled and react more harshly because the firm has failed to meet the high ethical standards it promotes.

These findings raise important questions about how CSR works when companies face tax-related and other financial misconduct. While earlier studies have focused on CSR's role in product failures, accidents, or general fraud, there is less evidence on how CSR affects market reactions to violations that directly weaken financial reporting and compliance. This study therefore investigates whether a strong CSR profile protects firms from negative stock market reactions when financial misconduct, such as tax violations, anti-money-laundering deficiencies, or accounting fraud, is revealed, or whether it can instead make reputational damage worse. By combining these related types of misconduct and applying an event study approach, this thesis aims to provide new evidence on CSR's role in this specific setting. The final sections of this introduction outline the research gap and the specific research question this study addresses.

1.1 Research gap

While prior research has examined how Corporate Social Responsibility (CSR) affects investor reactions to events such as product recalls, environmental incidents, and general fraud, much less is known about CSR's role in cases of financial misconduct that directly undermine reporting integrity and compliance. Studies like Godfrey et al. (2009) and Minor and Morgan (2011) suggest that CSR can build moral capital and protect firms during crises, yet findings by Bartov et al. (2019) show that this effect may reverse when stakeholders perceive misconduct as intentional deception. Despite this theoretical debate, the specific context of tax violations, anti-money-laundering failures, and accounting fraud remains under-researched, especially in terms of short-term abnormal returns. Existing studies often focus on single types of wrongdoing or treat CSR too broadly, without isolating its impact in settings where financial transparency is compromised. This study addresses this gap by combining three related types of financial misconduct into one framework and testing whether CSR truly mitigates or worsens reputational penalties when these issues come to light.

1.2 Research question and contribution

This thesis addresses the following central research question:

"To what extent does Corporate Social Responsibility (CSR) influence how financial markets react to announcements of financial misconduct?"

To answer this question, the study has three key objectives. First, it analyses whether announcements of financial misconduct lead to negative cumulative abnormal returns. Second, it tests whether a higher CSR score reduces or amplifies this market reaction. Third, it examines whether the severity of the misconduct influences the strength of the CSR effect. By combining a focused sample, a standard event study design, and a careful matching procedure, this thesis contributes new empirical evidence on CSR's role as a reputational shield, or risk factor, in the context of financial misconduct. Given the modest sample size and the small observed effect sizes, the results should be viewed as exploratory, providing a basis for further research on CSR and corporate reputation.

2 Literature review

2.1 Market reactions to tax violations

Understanding how financial markets react to corporate tax misconduct and related violations requires considering both short-term price movements and the broader reputational context. Corporate tax avoidance is often described as a continuum, ranging from legitimate tax planning through aggressive but legal schemes to outright evasion (Wang, Xu, & Sun, 2020). While tax planning within legal boundaries is accepted as sound financial management, aggressive tax strategies frequently push ethical limits, creating a risk that stakeholders may interpret them as intentional misconduct if revealed (Blaufus, Möhlmann & Schwäbe, 2019).

Empirical evidence supports this distinction. Blaufus et al. (2019) show that confirmed tax evasion leads to immediate and substantial declines in stock prices, whereas ordinary avoidance schemes typically do not harm valuation unless regulators or the media highlight them as improper. This pattern aligns with broader research on profit shifting and tax base erosion in multinational firms. Cooper and Nguyen (2020) explain how multinationals exploit internal transfer pricing, debt shifting, and intangible asset transfers to shift profits to low-tax jurisdictions. Beer, de Mooij, and Liu (2020) estimate that such activities cause significant tax revenue losses worldwide and frequently come under scrutiny through leaks and investigative journalism, as seen in the LuxLeaks and Panama Papers scandals. These events reveal how tax avoidance often overlaps with opaque structures, secrecy and other hints to corporate misconduct. Such activities often overlap with anti-money-laundering violations and accounting fraud when managers attempt to hide true financial flows.

The link between tax strategies and governance problems has been studied in depth. Kim, Li, and Zhang (2011) find that aggressive tax sheltering increases the risk of future stock price crashes because it gives managers greater discretion to hide bad news. When concealed information eventually comes to light, stock prices tend to drop sharply. Similarly, Hanlon and Slemrod (2009) show that market reactions to news about firms' involvement in tax shelters are generally negative, reflecting investor concerns about transparency and potential legal penalties.

Perception plays a critical role in shaping market outcomes. Ichev (2023) demonstrates that the tone and framing of media coverage can either magnify or mitigate investor reactions to misconduct. News stories that use strong negative language and clearly blame the firm tend to result in more severe stock price drops than neutral reporting. Davidson, Worrell, and Lee (1994) provide historical evidence that the market systematically penalizes firms involved in corporate crime, including tax evasion and bribery. Their findings also show that repeat offenders face stronger penalties, highlighting how reputational damage can accumulate over time.

Real-world scandals reinforce these findings. The LuxLeaks and Panama Papers exposed how well-known companies and wealthy individuals used complex offshore structures to avoid taxes and, in some cases, facilitate money laundering. Such revelations have intensified public scrutiny and regulatory crackdowns, making it increasingly risky for firms to rely on secrecy to protect aggressive tax practices. These examples illustrate why tax violations, anti-money-laundering deficiencies, and accounting fraud are interconnected forms of financial misconduct that share reputational consequences.

Taken together, these studies and events demonstrate that financial markets respond not only to direct legal consequences but also to broader concerns about management integrity, governance quality, and future cash flow uncertainty. This provides clear reasoning for analysing tax misconduct, anti-money-laundering failures, and accounting fraud within a combined framework.

2.2 The moderating effect of CSR

2.2.1 CSR as moral capital and reputational insurance

Corporate Social Responsibility (CSR) is often described as a resource that builds "moral capital" and serves as insurance against reputational damage when firms encounter negative events. Godfrey, Merrill, and Hansen (2009) propose that sustained CSR engagement earns goodwill from stakeholders, which can reduce unpleasant reactions when unexpected problems occur. Minor and Morgan (2011) find empirical support for this protective effect: firms with strong CSR track records experience smaller drops in stock prices following product recalls compared to firms with weaker CSR performance.

Shiu and Yang (2017) add further evidence from U.S. companies, showing that CSR investments are linked to less severe market penalties after controversial incidents. Their results suggest that consistent CSR activities can reassure investors that the firm remains committed to responsible conduct even when isolated failures occur. This aligns with stakeholder theory, which argues that CSR signals a broader commitment to multiple stakeholder groups, including communities, regulators, and investors.

The idea that CSR functions as reputational insurance is also relevant for financial misconduct contexts. When firms have high CSR credibility, stakeholders may perceive wrongdoing as accidental or isolated rather than systematic. This perception can mitigate immediate negative market reactions. Hanlon and Slemrod (2009) and Kim, Li, and Zhang (2011) indirectly support this by linking tax avoidance opacity to stock price crash risk: if CSR signals stronger internal controls and ethical culture, it might lower the likelihood that investors expect hidden risks.

2.2.2 CSR decoupling, hypocrisy, and backfire effects

Despite this protective view, a growing body of research shows that CSR does not always shield firms from reputational damage and may sometimes worsen it. Bartov, Marra and Momenté (2019) demonstrate that when firms issue financial restatements due to honest mistakes, a good CSR reputation reduces investor punishment. However, when restatements reveal intentional fraud, the same CSR reputation can backfire. Stakeholders feel misled by the gap between the firm's social claims and its unethical behaviour.

Ichev (2023) extends this to the media domain, finding that firms with high CSR scores face more severe public criticism when caught in wrongdoing, especially when misconduct appears to contradict stated values. This so-called hypocrisy penalty suggests that CSR can raise stakeholder expectations to a level that amplifies disappointment if a firm fails to meet its self-imposed standards.

Marques, Montenegro, and Brás (2025) provide an extensive meta-analysis that helps explain why the literature reports mixed findings on CSR and tax avoidance. They find that many firms engage in "CSR decoupling", promoting high-profile CSR initiatives publicly while continuing aggressive tax planning and other opportunistic behaviours behind the scenes.

Earlier work by Lanis and Richardson (2012) illustrates this dynamic in the tax context: firms may increase CSR disclosures to offset reputational risks tied to aggressive tax strategies. Hoi, Wu, and Zhang (2013) show that firms with irresponsible CSR profiles are statistically more likely to shelter taxes, reinforcing the link between poor ethical culture and opportunistic fiscal behaviour. Desai and Dharmapala (2006) add a theoretical dimension by describing how tax avoidance can be used by managers to extract private benefits, intensifying agency costs and undermining the credibility of CSR signals.

Mkadmi and Ben Ali (2024) confirm this tension in emerging markets: in Tunisia's banking sector, CSR participation discourages tax avoidance behaviour overall, but firms with high profitability and strong sales incentives sometimes pursue aggressive tax strategies despite their CSR commitments. This mixed evidence highlights that CSR's effectiveness as a shield depends heavily on whether stakeholders believe that a firm's actions match its social promises.

In summary, these studies indicate that CSR can have two sides: it may cushion the reputational impact of financial misconduct when seen as authentic and consistent, but it can also intensify backlash when stakeholders perceive a disconnect between what the firm says and what it does. This tension is central to this thesis, which tests whether CSR moderates investor reactions to financial misconduct involving tax violations, anti-money-laundering failures, and accounting fraud.

2.3 Reputation, stakeholder theory, and agency costs

The mixed evidence on CSR's moderating effect is grounded in several theoretical perspectives that help explain how and why stakeholders react to financial misconduct. Stakeholder theory argues that firms have responsibilities not only to shareholders but also to a wider group, including regulators, communities, and the public. When firms engage in misconduct, these groups may respond through negative publicity, consumer boycotts, or political pressure, amplifying reputational costs beyond immediate financial penalties.

Agency theory provides another perspective, which shows how managers may pursue tax avoidance and profit shifting to maximise short-term gains, even when this conflicts with shareholder interests (Desai & Dharmapala, 2006; Kim et al., 2011). Wang et al. (2020) emphasize that greater managerial discretion in tax planning can increase information asymmetry and moral hazard, making it more likely that managers will exploit opacity to conceal risks. Such behaviour raises the probability of abrupt corrections in firm value when hidden misconduct is exposed.

Signalling theory suggests that CSR activities send signals to stakeholders about a firm's ethical standards and commitment to long-term sustainability. However, when actual behaviour contradicts these signals, the inconsistency can damage credibility more severely than if the firm had no CSR profile at all (Bartov et al., 2019; Ichev, 2023). This interplay between agency costs, stakeholder expectations, and signalling forms the theoretical basis for examining whether CSR can effectively buffer reputational damage from financial misconduct or whether it may sometimes backfire.

2.4 Summary and research gap

In summary, prior research shows that financial markets react negatively to revelations of tax violations, anti-money-laundering failures, and accounting fraud due to concerns about governance quality and future risk. CSR can moderate these reactions by building reputational goodwill but may also intensify penalties when stakeholders perceive inconsistency between CSR claims and actual misconduct. Although recent meta-analyses and empirical studies provide valuable insights, there remains a need for event-based evidence specifically linking CSR, violation severity, and investor responses within the context of financial misconduct that threatens reporting transparency. This study addresses this gap by testing whether CSR serves as a protective buffer or a reputational risk amplifier when firms face tax-related and related financial controversies.

3 Hypothesis development

Building on the theoretical framework and prior literature discussed in Chapter 2, this chapter outlines the hypotheses that guide the empirical analysis of this thesis. These hypotheses address the main research question:

"To what extent does Corporate Social Responsibility (CSR) influence how financial markets react to announcements of financial misconduct?"

This chapter is structured to present each hypothesis in order:

H1: The baseline effect of financial misconduct announcements on stock returns.
H2: The moderating effect of CSR.
H3: How this moderation varies with the severity of the violation.

3.1 Financial misconduct and short-term market reaction

The first hypothesis builds on extensive event study research showing that corporate misconduct typically results in negative abnormal stock returns (Davidson et al., 1994; Kim, Li, & Zhang, 2011). When firms are found to have violated tax laws, anti-money-laundering regulations, or accounting standards, investors often perceive this as a signal of weak governance, heightened regulatory risk, and potential future costs (Hanlon & Slemrod, 2009; Cooper & Nguyen, 2020).

Signalling theory suggests that such disclosures reveal hidden risks that were previously obscured by opaque reporting. Agency theory also implies that managers may have used aggressive financial practices for private benefit, harming shareholder value when exposed. Empirical studies on tax shelters, profit shifting, and fraud provide consistent evidence that these announcements trigger short-term price declines.

Hypothesis 1 (H1)

Announcements of tax violations, anti-money-laundering deficiencies, or accounting fraud are associated with negative abnormal stock returns.

3.2 The moderating role of Corporate Social Responsibility

The second hypothesis extends this baseline by examining whether CSR moderates the market reaction. As discussed in Chapter 2, the moral capital perspective (Godfrey et al., 2009) argues that firms with strong CSR profiles enjoy greater stakeholder trust and reputational goodwill, which can cushion the impact of negative events. Stakeholder theory reinforces this idea by suggesting that high-CSR firms are perceived as more ethical and responsible, making misconduct seem less intentional or systemic.

However, as noted in Bartov et al. (2019) and Ichev (2023), CSR can also backfire if stakeholders see misconduct as contradicting the firm's proclaimed values. This "hypocrisy penalty" means that CSR might protect reputation only when stakeholders believe it is genuine.

Despite this mixed theoretical expectation, most studies find an average buffering effect for moderate misconduct. Therefore, the second hypothesis tests whether strong CSR reduces the level of negative abnormal returns when misconduct is revealed.

Hypothesis 2 (H2)

Firms with a high CSR profile experience a less negative market reaction to announcements of financial misconduct compared to firms with a low or average CSR profile.

3.3 Conditional moderation based on severity

The third hypothesis explores whether the reputational benefit of CSR depends on how severe and serious the financial misconduct is perceived to be. Severity is captured through two complementary measures: the monetary penalty imposed (log-transformed) and whether the violation was handled as a civil or criminal case. Higher penalties and criminal classifications typically signal greater misconduct seriousness and attract stronger public and regulatory scrutiny.

The underlying expectation is that while CSR may provide reputational insurance for less severe or ambiguous misconduct, this protective effect may decrease, or even reverse, when stakeholders perceive a serious breach of trust. Evidence by Marques et al. (2025) suggests that severe wrongdoing exposes the gap between CSR narratives and actual behaviour, leading to greater stakeholder disappointment and more intense market penalties. From a signalling theory perspective, severe or criminal violations undermine the credibility of a firm's social responsibility image, intensifying negative reactions when misconduct comes to light.

Therefore, the third hypothesis tests if the moderating effect of CSR becomes weaker, or may even backfire, as the severity and seriousness of the violation increase.

Hypothesis 3 (H3)

The moderating effect of CSR on the market reaction to financial misconduct weakens or reverses as the severity of the violation increases, as indicated by higher monetary penalties and criminal prosecution.

4 Methodology

This research follows a quantitative and event-based approach to examine how financial markets respond to announcements of corporate financial misconduct and how Corporate Social Responsibility (CSR) moderates this reaction. Specifically, it tests whether short-term stock price reactions differ for firms with stronger CSR profiles and whether this moderating effect depends on the perceived severity of the violation. This study conceptualizes CSR in line with stakeholder theory and measures it quantitatively using the Refinitiv ESG Combined Score as an objective indicator of a firm's environmental, social, and governance performance.

The focus is on publicly listed companies between 2010 and 2024, a period selected to reflect increased regulatory scrutiny and the growing importance of CSR in investment decision-making. Event data are drawn from the Good Jobs First Violation Tracker, while ESG and market data are sourced from Refinitiv (LSEG). Financial firm data is retrieved from Compustat.

4.1 Independent, dependent and control variables

The main independent variable, [TaxViolation], is a binary indicator coded 1 if a firm has a publicly disclosed violation related to tax evasion, anti-money-laundering deficiencies, or accounting fraud during the sample period. This broader definition captures different forms of financial misconduct that similarly weaken financial transparency and signal governance risk, aligning with prior research (Davidson et al., 1994; Kim, Li, & Zhang, 2011; Ichev, 2023). The event date is defined as the first public announcement of the violation.

The main dependent variables in this study are the cumulative abnormal returns (CAR) over the short event windows of [-1, +1] and [-3, +3] trading days (CAR_pm1 and CAR_pm3). These values measure the market's reaction to tax violation announcements. Cumulative abnormal returns are estimated using the market model and verified with the Fama-French three-factor model to control for market-wide risk factors. Simpler daily abnormal returns are not analysed separately, as they are implicitly captured within the CARs.

In all models, standard firm-level controls are included to account for alternative explanations of abnormal stock returns. These controls include firm size (log of total assets), leverage (debt-to-assets), return on assets (ROA) and industry. These variables are sourced from Refinitiv and Compustat and merged into the event dataset using ISIN and CIK identifiers.

4.2 Sample selection and data collection

4.2.1 Firm matching variables

To ensure that the estimated effect of financial misconduct on stock market reactions is not driven by systematic differences between treated and untreated firms, this study applies a matched-sample design. Matching aims to create a control group of firms that are similar to the treated firms (those with announced violations) in key observable characteristics, thus improving the reliability of causal inferences.

For each firm with a confirmed tax violation, anti-money-laundering deficiency, or accounting fraud disclosure, at least one comparable control firm was identified using propensity score matching with nearest neighbour selection. Matching was based on firm-level financial and structural variables known to influence stock returns and corporate misconduct risk.

- Firm size

Log of Total Assets (log_at): firm size was operationalized as the natural logarithm of total assets rather than using raw total assets. Taking the logarithm reduces skewness and ensures comparability between large and small firms, aligning with standard practice in corporate finance and event studies. Larger firms typically have more resources and may attract different levels of regulatory scrutiny and media attention than smaller firms.

- Profitability

Return on Assets (ROA): captures operational efficiency and indicates how effectively a firm turns its assets into profits. Firms with higher ROA may have less incentive for aggressive tax strategies or financial manipulation.

Net Income (ni): included to reflect overall bottom-line financial performance. Unlike total assets, net income was matched in its raw form rather than as a logarithm. This choice was made because net income can be negative for loss-making firms, making log transformation technically infeasible. Given this limitation, the raw value was used as a practical trade-off to ensure that both profitable and unprofitable firms could be meaningfully matched.

- Leverage

Leverage Ratio: captures the proportion of debt financing relative to total assets. Highly leveraged firms may face greater pressure to engage in earnings management or aggressive tax planning.

- Industry

Industry Code (SIC): controls for industry-specific risk factors and reporting norms. Firms were matched within the same industry classification (4-digits) to account for sector effects on misconduct likelihood and investor reactions.

All matching variables were lagged by one year relative to the event date to minimize bias from potential changes in firm characteristics immediately preceding the disclosure of misconduct.

By matching on these variables, the analysis ensures that differences in stock price reactions can be more confidently attributed to the announcement of financial misconduct and the moderating role of CSR, rather than to underlying differences in firm size, profitability, capital structure, or industry context.

The matching procedure was carried out using the MatchIt package in R with a nearest neighbour matching with a 1:1 to 1:5 ratio, depending on available controls per treated firm, was applied. This matching procedure was chosen to identify the most similar control firms for each treated firm based on the chosen financial and industry variables. Due to data availability and the distribution of firms within subclasses, the number of control firms per treated firm varies within the dataset: some treated firms have only one closely comparable control, while others have up to three, four, or in a few cases five matched controls. This variation was necessary to maximise matching quality and ensure that each treated firm had at least one meaningful counterpart. While this approach enhances balance overall, it also means that some observations rely on a narrower comparison base than others. This is taken into account when interpreting the regression results and conducting robustness checks.

4.2.2 Stock data and abnormal returns

To measure how financial markets respond to announcements of tax violations, anti-money-laundering deficiencies, or accounting fraud, this study relies on cumulative abnormal returns (CARs) around each event date. The required daily stock price data were collected from Refinitiv (LSEG Workspace), which provides high-quality, historical price series for publicly listed firms. For each firm in the sample, the event date was defined as the official announcement date of the violation. Using this date, daily closing prices were extracted for a window starting 250 trading days before the event date up to at least 10 trading days after the announcement. This reference period ensures a sufficiently long estimation window for computing normal expected returns. Daily stock returns were calculated as the simple percentage change between consecutive closing prices.

To adjust for general market movements and risk factors, abnormal returns were estimated using the Fama-French three-factor model. Daily values for the market risk premium, SMB (size factor), and HML (value factor) were retrieved from Kenneth French's data library and matched to the same time period as the firm returns. This procedure ensures that the CARs reflect market reactions attributable specifically to the disclosure of financial misconduct, adjusted for normal market risk and firm-specific characteristics. The estimation window for the model parameters covered trading day -250 to trading day -11 relative to the event date, following standard practice to avoid contamination by information leaks close to the event. The model for expected returns was specified as:

$$R_{it} - R_{ft} = \alpha_i + \beta_{MKT}(R_{Mt} - R_{ft}) + \beta_{SMB} SMB_t + \beta_{HML} HML_t + \epsilon_{it}$$

where:

- R_{it} is the stock return for firm I on day t,
- R_{ft} is the risk-free rate on day t,
- R_{Mt} is the return on the market portfolio,
- β_{MKT} is the sensitivity of the fir m's return to the market return
- SMB_t and HML_t are the size and value factors respectively.

Based on these expected returns, abnormal returns (ARs) were computed as the difference between actual and expected returns. Cumulative Abnormal Returns (CARs) are then obtained by summing ARs across specific event windows. In practice, CARs were computed directly in R using, but they reflect the theoretical sum of daily abnormal returns over the event period. Specifically, CARs were calculated for windows of [-1,+1], [-3,+3], [-5,+5] and [-10,+10] trading days surrounding the announcement date.

Full sets of CARs and ARs for all event windows are included in the underlying dataset. However, only [-1,+1] and [-3,+3] CARs are presented in the main regression results, as they are standard in the literature and offer the clearest insight into short-term market responses. The symmetric windows allow detection of possible information leakage, while the post-event windows help assess the persistence or correction of the market reaction after the fine is disclosed.

4.2.3 ESG data

The ESG data for this study were obtained from Refinitiv, using the ESG Combined Score as the primary measure. This score provides a comprehensive indicator of a firm's performance on environmental, social, and governance issues while also factoring in any controversies that could affect its reputation. It is derived through a transparent, rules-based system that combines results from ten underlying categories, including areas like carbon emissions, labour practices, and board oversight.

The Combined Score merges the standard ESG score with a controversy layer, which adjusts the rating to account for negative incidents or stakeholder disputes. This means the score reflects not just what companies report about their ESG activities, but also how they are perceived in terms of risks and reputational challenges. To avoid bias from post-event adjustments, all ESG scores were lagged by one year relative to the event date.

All scores are calculated based solely on publicly available information, such as annual reports, CSR statements, and official filings with regulators. Refinitiv's methodology avoids using estimated or modelled data, ensuring that the ratings rely on actual, verifiable disclosures that can be compared consistently across firms and over time.

The final ESG scores are standardized on a 0 to 1 scale, where higher values indicate stronger CSR performance. To explore potential non-linear effects, a dummy variable (HighCSR) was created for firms in the top 25% of ESG Combined scores. This allows comparison between high-CSR firms and all others (middle and low ESG performers). The regressions test whether firms with distinctly strong CSR reputations experience different market reactions to misconduct.

Although the ESG Combined Score is a widely accepted measure in ESG research, it is worth noting that differences in scoring methods across data providers remain an important consideration. Refinitiv's approach places strong emphasis on the quantity and consistency of publicly disclosed information, which can lead to differences compared to more qualitative or analyst-driven ratings from other sources like MSCI or Sustainalytics.

4.2.4 Violation severity variables

In addition to firm matching and core financial variables, this study includes detailed measures to capture the seriousness and legal nature of each financial misconduct case. These variables provide a more precise basis for testing whether the effectiveness of Corporate Social Responsibility (CSR) as reputational insurance depends on how severe a violation is perceived to be (H3). Due to the strongly right-skewed nature of penalty data, with a few exceptionally large fines, the raw values were transformed using the natural logarithm. This log transformation (logPenalty) standardizes the scale and minimizes the influence of outliers, in line with established empirical research on regulatory sanctions (e.g., Davidson et al., 1994; Bartov et al., 2019). No violations are disclosed without an explicit monetary penalty, so undisclosed data did not have to be taken into consideration.

In addition to financial severity, the legal nature of each violation was recorded to distinguish between civil and criminal cases. This distinction acknowledges that criminal cases often carry stronger legal and reputational consequences than civil infractions, which may influence how investors interpret the event and the effectiveness of CSR as a reputational buffer. The information was stored as a binary variable (civil_criminal), with 0 indicating a civil violation and 1 indicating a criminal violation. This variable is included as a control in the third regression model, allowing the analysis to test whether CSR moderates investor responses differently depending on the seriousness of the misconduct.

4.3 Data merging and cleaning

The data merging and cleaning process was carried out in multiple phases to construct a high-quality sample containing only relevant and comparable firms. The starting point was the Good Jobs First Violation Tracker database, which documents corporate misconduct cases globally. For this study, the Global Tracker, the US Tracker, and the UK Tracker were first combined into a single working file.

From this combined dataset, records were filtered to include only cases where the primary offense was identified as a tax violation, money laundering, or accounting fraud or deficiencies, in line with the scope of the research. To ensure relevance to current corporate governance standards and the availability of sufficient valuable ESG Combined Scores, only violations occurring after 2010 were retained. Additionally, only publicly traded companies were included to ensure the availability of market and financial data.

Duplicate entries were then identified and removed, as some violations appeared multiple times due to the involvement of several regulatory agencies in the same case. This step ensured that each misconduct event was represented only once in the final dataset.

To enable the integration of financial and market data, the violation records were matched to Compustat identifiers. For global firms, ISINs were generally available and used directly for matching. For US firms lacking an ISIN, CIK codes were used to retrieve the gvkey, Compustat's unique firm-level identifier. Using the gvkey, the corresponding stock tickers and CUSIPs were obtained, as Refinitiv requires tickers, exchange codes (RICs), or ISINs for data extraction rather than CIK or gvkey.

After collecting all necessary identifiers, the required matching variables, including log of total assets, ROA, leverage, net income, and SIC code, were retrieved from Compustat. Based on these variables, a matched sample of suitable control firms was constructed. Initially, this process provided at least three appropriate controls per violation, resulting in 150 valid violation cases and a total of 701 control firms, yielding 851 firm-event observations overall. For each control firm, a pseudo-event date was assigned based on the violation date of its matched treated firm. Violation cases were coded as treated = 1 and control firms as treated = 0.

Due to incomplete identifiers and missing gvkeys for a large amount of control firms, the pool of available controls decreased significantly during this matching phase. After applying strict matching criteria and excluding observations lacking essential data, the control group was reduced to 322 firms. As a result, some violations remained with only one matched control firm instead of the intended minimum of three. The decision was made to maintain strict matching thresholds rather than expand the pool with less comparable controls, prioritizing the quality of the matching over quantity.

Following the removal of remaining missing data, the refined dataset contained 408 observations, including 127 violation cases. This dataset was then matched with daily stock prices from Refinitiv and ESG scores from Refinitiv's ESG database. To compute the cumulative abnormal returns, each firm was required to have complete daily stock price data for both a reference window (approximately 250 trading days prior to the event date) and an event window (−10 to +10 trading days around the event date). During this final stage, all observations without proper stock data or a viable ESG Combined Score were removed, after which a sample of 80 treated firms and 157 matched control firms remained.

4.4 Final sample

The final sample consists of publicly listed firm-year observations between 2010 and 2025. Observations missing key variables or identifiers were excluded. Only records with complete data for the dependent and control variables were retained. The cleaned and merged dataset, saved as final_thesis_dataset.csv, forms the basis of the empirical analysis in chapter 5.

4.5 Empirical models

To empirically test the hypotheses presented in chapter 3, the following regression models were constructed. Each model incorporates lagged variables of 1 financial year.

4.5.1 *Model 1: main effect of financial misconduct (H1)*

Model 1 estimates whether announcements of tax violations, AML deficiencies, or accounting fraud result in negative abnormal returns.

$$CAR_{it} = \beta_0 + \beta_1 \cdot \text{TaxViolation}_{it} + \beta_2 \cdot \text{Controls}_{it} + \varepsilon_{it}$$

Here, CAR_{it} is the cumulative abnormal return for firm *i* during the event window. A negative and significant β_1 supports H1, indicating that markets react negatively when such misconduct becomes public.

4.5.2 *Model 2: moderating role of CSR (H2)*

The second model adds an interaction term to test whether a strong CSR profile buffers or amplifies the market penalty associated with financial misconduct.

$$\begin{aligned} CAR_{it} &= \beta_0 + \beta_1 \cdot \text{TaxViolation}_{it} + \beta_2 \cdot \text{HighCSR}_{it} \\ &+ \beta_3 \cdot (\text{TaxViolation}_{it} \times \text{HighCSR}_{it}) + \beta_4 \cdot \text{Controls}_{it} + \varepsilon_{it} \end{aligned}$$

A positive and significant β_3 would indicate that firms with high CSR scores experience less negative abnormal returns, consistent with the reputational insurance hypothesis in H2.

4.5.3 *Model 3: conditional moderation based on severity (H3)*

In the third regression model, severity is proxied by both the log-transformed penalty amount and an indicator for criminal cases (civil_criminal). An interaction term between HighCSR and each severity proxy tests whether CSR's effect varies with seriousness of the violation. This allows the analysis to distinguish whether high CSR still buffers negative market reactions when misconduct is severe or criminal in nature.

$$CAR_{it} = \beta_0 + \beta_1 HighCSR_{it} + \beta_2 logPenalty_{it} + \beta_3 CivilCriminal_{it} + \beta_4(HighCSR \times logPenalty) + \beta_5(HighCSR \times CivilCriminal) + \text{controls} + \epsilon_{it}$$

A positive and significant coefficient on the interaction term HighCSR:log_penalty_adjusted would suggest that CSR continues to provide some reputational cushion even when penalties are high, contrary to the expectation that severe misconduct should erode CSR's protective effect. A similar interpretation applies for HighCSR:Civil_Criminal_Numeric with respect to criminal cases.

4.6 Estimation approach and robustness checks

All regressions were estimated using ordinary least squares (OLS). While robust standard errors are often used to account for heteroskedasticity, the models in this study do not explicitly apply them. This is noted as a limitation. Year and industry fixed effects are included in all models to control for unobserved time and sector-specific influences. Robustness checks include:

- Testing alternative CAR windows: [-1,+1], [-3,+3], [-5,+5] and [-10,+10]
- Using both the ESG Score and ESG Combined Score to measure CSR
- Testing subsamples for civil versus criminal cases
- Post-hoc power analyses to assess whether the sample size is sufficient to detect the hypothesised effects.

Alternative CAR windows produced qualitatively similar results, though significance levels varied slightly. Results using the ESG (standard) score instead of the Combined Score yielded consistent patterns. Subsample analysis showed no systematic difference in CSR effects between civil and criminal cases. Post-hoc power tests confirmed limited ability to detect interaction effects, particularly in H2 and H3.

4.7 Statistical power and limitations

While the matched sample design and standardized event study approach strengthen the reliability of the findings, certain limitations remain. First, the sample size is constrained by the availability of financial data and verified CSR scores, which may limit statistical power for detecting small moderation effects. Given the final sample size of 237 firm-event observations, with 80 treated firms and varying numbers of matched controls, statistical power is limited for detecting small to moderate effects, especially for interaction terms involving ESG scores and severity measures.

A formal post hoc power analysis was conducted to assess the sensitivity of the dataset. For Hypothesis 1 (H1), which tests the average market reaction to announcements of financial misconduct using a difference-in-means framework, the power analysis was based on a two-sided t-test comparing 80 treated firms and 157 matched controls. The calculated effect size (Cohen's d) was approximately 0.0047, resulting in a statistical power of just 0.051 at the 5% significance level. This suggests that the data is extremely underpowered for detecting small treatment effects.

Similarly, for Hypotheses 2 and 3 (H2 and H3), which involve multiple regression models testing for moderation effects of CSR and severity, the power analysis was conducted for a model with one tested predictor (the interaction term) and 230 degrees of freedom. The observed effect size (Cohen's f^2) was approximately 0.00116, yielding an even lower power of 0.081. This confirms that the study is substantially underpowered for detecting interaction effects in the regression framework.

These results highlight a critical limitation: while the estimated models are theoretically motivated and statistically well-specified, the sample is not large enough to reliably detect the subtle moderation effects hypothesized in H2 and H3. As a result, the absence of statistically significant findings in these models should not be interpreted as definitive evidence of no effect, but rather as an indication of insufficient statistical power.

Another methodological limitation is that the regression analyses were conducted using classical OLS standard errors without robust correction for potential heteroskedasticity. While OLS estimates remain unbiased, the standard errors, and thus significance levels, may be affected if error variance is not constant across observations.

Lastly, CSR measures rely on self-reported or third-party compiled data, which may not fully capture how stakeholders perceive a firm's ethical conduct, particularly in light of potential CSR decoupling. These factors should be considered when interpreting the results and generalizing the conclusions.

5 Results and empirical analysis

This chapter presents the results from the regression analyses used to test the three hypotheses introduced in chapter 3. It starts with descriptive statistics and correlation results, followed by the regression outcomes for each hypothesis. The main focus is on H2, which tests the association between CSR performance and the stock market reaction on announcements of financial misconduct.

Table 1 provides an overview of the key variables used in the empirical analysis. The final sample consists of 237 firm-event observations, including 80 recorded violation cases and 157 matched control firms.

5.1 Descriptive analysis

Table 1. Descriptive Statistics

Variable	Obs	Mean	SD	Min	Max
CAR_pm1	237	-0.002565154	0.04250342	-0.1793438	0.3829345
CAR_pm3	237	-0.007443478	0.07194727	-0.2451974	0.5627394
CAR_pm5	237	-0.009813206	0.08153958	-0.2724159	0.5542669
CAR_pm10	237	0.016747687	0.26068660	-0.4011891	2.5185538
ESGC_Score	237	54.242572740	16.83106957	12.1112050	87.8950912
log_penalty_ adjusted	80	15.771388002	3.35331569	8.5943394	21.4164130
Civil_Criminal_ Numeric	80	0.200000000	0.40252368	0.0000000	1.0000000
log_at	237	11.895304432	2.37482021	5.8529436	15.1856334
roa	237	0.031557102	0.06148422	-0.1452240	0.5646411
leverage	237	0.222899744	0.15744490	0.0000000	0.8795613
ni	237	4,305.098751055	5,768.25014153	-10,938.3610000	25,794.0000000

The mean value of the short-term cumulative abnormal return for the window of [-1, +1] days (CAR_pm1) is approximately -0.0026, with a minimum of -0.1793 and a maximum of 0.3829. This suggests that, on average, firms in the sample experienced a slight decline in stock returns during the immediate period around the violation or pseudo-event announcement date, though the range indicates substantial variation across individual cases.

For longer windows, the mean cumulative abnormal returns are slightly more negative. For example, CAR_pm3 ([-3, +3] days) has a mean of about -0.0074, and CAR_pm5 ([-5, +5] days) averages -0.0098. The largest event window considered, CAR_pm10 ([-10, +10] days), shows a small positive mean of 0.0167, although the minimum and maximum values indicate considerable spread and outlier effects.

The ESG Combined Score (ESGC_Score), which quantifies each firm's overall environmental, social, and governance performance including controversies, has an average of 54.24 in the sample. Scores range from as low as 12.11 to as high as 87.90, indicating substantial variation in sustainability and governance practices across firms.

Regarding the severity of misconduct, the natural logarithm of the adjusted penalty (log_penalty_adjusted) has a mean of approximately 15.77, with a minimum of 8.59 and a maximum of 21.42. This reflects the wide distribution of monetary penalties imposed on firms in the sample.

The civil_criminal variable indicates whether a violation was classified as civil, criminal, or a combination of both. This categorical classification adds context to the severity and legal implications of each misconduct event.

The firm-level control variables show expected variation. The average log of total assets (log_at) is 11.90, indicating that the sample consists primarily of medium to large firms. Profitability (roa) averages around 0.032, which is reasonable for a diverse sample of public companies. The mean leverage ratio (leverage) is about 0.223, with values ranging up to nearly 0.88, reflecting a broad range of capital structures. Net income (ni) shows considerable spread, with a mean of 4,305 and values ranging from large losses to substantial profits.

The final sample covers a range of firm profiles, but the relatively limited number of events and matched controls means that estimates should be interpreted with caution, particularly for interaction effects and subgroup analyses.

5.2 Correlation Matrix

Table 2: Correlation Matrix of Key Variables

	CAR_pm1	CAR_pm3	CAR_pm5	CAR_pm10	ESGC_Score	log_penalty_adjusted	log_at	roa	leverage	ni	Civil_Criminal_Numeric
CAR_pm1	1.00										
CAR_pm3	0.69	1.00									
CAR_pm5	0.63	0.91	1.00								
CAR_pm10	0.28	0.40	0.40	1.00							
ESGC_Score	0.03	-0.04	-0.04	0.03	1.00						
log_penalty_adjusted	-0.12	-0.30	-0.29	-0.38	-0.11	1.00					
log_at	-0.19	-0.25	-0.20	-0.18	0.12	0.50	1.00				
roa	0.43	0.19	0.18	0.15	-0.06	-0.21	-0.52	1.00			
leverage	-0.07	0.02	0.06	-0.01	0.07	0.01	-0.13	-0.14	1.00		
ni	-0.03	-0.01	0.01	-0.03	0.04	0.20	0.55	-0.08	-0.08	1.00	
Civil_Criminal_Numeric	-0.19	-0.15	-0.14	-0.15	-0.02	0.49	0.35	-0.16	-0.25	0.14	1.00

-1 -0.75 -0.5 -0.25 0 0.25 0.5 0.75 1

Table 2 presents the pairwise Pearson correlation coefficients for the main variables used in the empirical analysis. As expected, the different cumulative abnormal return (CAR) measures are positively correlated with each other. The strongest correlations are found between the shorter event windows: CAR_pm3 and CAR_pm5 show a very high correlation (0.91), indicating that stock market reactions are fairly consistent across overlapping short-term windows. The correlation between CAR_pm1 and the longer CAR_pm10 is noticeably weaker (0.28), which aligns with the idea that immediate market reactions may differ from patterns over a longer horizon as new information is absorbed.

The ESG Combined Score shows very low correlations with both the CARs and the other firm-level financial variables, suggesting that ESG performance, as measured here, is not strongly related to short-term abnormal returns or to basic financial structure. This low correlation supports the interpretation that the ESG variable captures an independent dimension of a firm's profile rather than overlapping with size or profitability.

Penalty severity, captured by the log-transformed penalty amount, is moderately negatively correlated with the CARs, especially for longer windows (for example -0.38 with CAR_pm10). This suggests that more severe violations tend to be associated with more negative market reactions. Firm size, measured as the log of total assets, is moderately positively correlated with the penalty amount (0.50), which is plausible since larger firms may face higher absolute fines when misconduct occurs.

The highest correlation between controls is between firm size and net income (0.55), which is expected since larger firms generally report higher profits. Return on assets (ROA) shows a moderate negative correlation with firm size (–0.52), reflecting the well-known pattern that very large firms sometimes operate with lower asset efficiency than smaller, more agile companies.

The pairwise Pearson correlation matrix suggests that the independent and control variables are not excessively correlated, as no coefficients approach levels typically associated with multicollinearity. While this supports the assumption of distinct underlying constructs, formal multicollinearity diagnostics (for example VIF) were not conducted. Therefore, the regression results should be interpreted with caution regarding potential overlap in explanatory power among variables.

5.3 Mean Difference Tests

Table 3: Mean Difference Tests for CARs

CAR Window	Mean (Control)	Mean (Treated)	t-Statistic	p-Value	95% CI Lower	95% CI Upper
CAR_pm1	−0.0026	−0.0024	−0.0274	0.9782	−0.0140	0.0136
CAR_pm3	−0.0106	−0.0013	−0.7700	0.4431	−0.0331	0.0146
CAR_pm5	−0.0114	−0.0067	−0.3524	0.7253	−0.0314	0.0219
CAR_pm10	−0.0015	0.0525	−1.1574	0.2502	−0.1465	0.0387
CAR_post1	−0.0024	−0.0077	0.6987	0.4864	−0.0098	0.0204
CAR_post3	−0.0024	−0.0033	0.1044	0.9170	−0.0167	0.0185
CAR_post5	−0.0011	−0.0035	0.2203	0.8261	−0.0192	0.0240
CAR_post10	0.0102	−0.0004	0.1657	0.8688	−0.1159	0.1370

To examine whether firms that announced financial misconduct experience significantly different short-term stock market reactions than comparable non-violating firms, independent sample t-tests were conducted for each cumulative abnormal return (CAR) window. Both symmetric event windows (which capture possible information leakage and immediate market adjustment) and purely post-announcement windows were tested.

Table 3 summarizes the results. For the symmetric windows, the average CARs for treated firms (violations) are slightly less negative than those for control firms for the shortest window (CAR_pm1), but this difference is extremely small and statistically insignificant ($t = -0.027$, $p = 0.978$). Similarly, for CAR_pm3 and CAR_pm5, the mean differences remain small and are not statistically significant at conventional levels (p-values of 0.44 and 0.73, respectively). For the widest symmetric window (CAR_pm10), the mean CAR for treated firms is slightly positive on average (0.052), while for control firms it remains slightly negative (-0.001), but again this difference is not statistically significant ($p = 0.25$).

For the purely post-event windows, which isolate the reaction in the days following the announcement, the results show a consistent pattern of minor differences without statistical significance. For example, the mean abnormal return for day +1 (CAR_post1) is slightly more negative for treated firms than controls, but the difference is not significant ($p = 0.49$). The same holds for longer post-event windows: CAR_post3, CAR_post5 and CAR_post10 all show very small and statistically insignificant mean differences, with p-values well above the 0.05 threshold.

Overall, these results suggest that the mean abnormal returns do not differ significantly between firms with announced violations and their matched controls across both symmetric and purely post-announcement windows. This finding implies that, on average, the announcement of tax, money laundering, or accounting misconduct does not produce a statistically significant immediate stock price reaction in this sample. However, as mean difference tests alone do not control for other firm characteristics or variation in violation severity, more detailed regression analyses are required to test the hypotheses comprehensively.

5.4 Regression results of hypothesis 1

For each hypothesis, regression results are reported for two event windows: a short window (CAR_pm1) capturing the immediate stock market reaction, and a slightly longer window (CAR_pm3) capturing any delayed adjustment. Together, these results shed light on how financial markets respond to corporate misconduct and the moderating role of CSR, while also accounting for severity and legal nature. We first discuss hypothesis 1.

Table 4: Regression Results for H1 (CAR_pm1)

Variable	Estimate	Std. Error	t-Statistic	p-Value
(Intercept)	−0.0253	0.0199	−1.2669	0.2065
treated	−0.0036	0.0054	−0.6691	0.5041
log_at	0.0012	0.0016	0.7450	0.4571
roa	0.3234	0.0520	6.2177	0.0000
leverage	0.0025	0.0169	0.1480	0.8825
ni	0.0000	0.0000	−0.3713	0.7107

Table 5: Regression Results for H1 (CAR_pm3)

Variable	Estimate	Std. Error	t-Statistic	p-Value
(Intercept)	0.0961	0.0358	2.6813	0.0079
treated	0.0076	0.0098	0.7816	0.4353
log_at	−0.0097	0.0028	−3.4004	0.0008
roa	0.0370	0.0935	0.3954	0.6929
leverage	−0.0051	0.0304	−0.1687	0.8662
ni	0.0000	0.0000	2.1182	0.0352

Table 4 and Table 5 display the results for the first hypothesis, which tests whether the announcement of financial misconduct (tax violations, anti-money laundering deficiencies, or accounting fraud) is associated with abnormal stock returns.

For the immediate window (CAR_pm1), the coefficient for treated is negative but not statistically significant ($p = 0.5041$). This suggests that, on average, announcements do not trigger a clear immediate abnormal return within the [-1, +1] window.

For the slightly longer window (CAR_pm3), the coefficient for treated is positive and also not significant ($p = 0.4353$). This implies that no consistent market penalty or reward is observed in the days following the announcement either.

Taken together, the evidence does not provide strong support for H1. Investors do not appear to react systematically to the disclosure of financial misconduct, at least in aggregate, within short-term event windows.

5.5 Regression results of hypothesis 2

Table 6: Regression Results for H2 (CAR_pm1)

Variable	Estimate	Std. Error	t-Statistic	p-Value
(Intercept)	−0.0255	0.0201	−1.2697	0.2055
treated	−0.0027	0.0062	−0.4362	0.6631
HighCSR	0.0037	0.0070	0.5295	0.5970
log_at	0.0011	0.0016	0.7051	0.4814
roa	0.3244	0.0522	6.2086	0.0000
leverage	0.0019	0.0170	0.1106	0.9120
ni	0.0000	0.0000	−0.3951	0.6932
treated:HighCSR	−0.0031	0.0133	−0.2332	0.8158

Table 7: Regression Results for H2 (CAR_pm3)

Variable	Estimate	Std. Error	t-Statistic	p-Value
(Intercept)	0.1011	0.0360	2.8085	0.0054
treated	0.0155	0.0111	1.3929	0.1650
HighCSR	0.0081	0.0126	0.6436	0.5205
log_at	−0.0103	0.0029	−3.5764	0.0004
roa	0.0413	0.0935	0.4419	0.6590
leverage	−0.0060	0.0305	−0.1982	0.8430
ni	0.0000	0.0000	2.0507	0.0414
treated:HighCSR	−0.0366	0.0238	−1.5399	0.1250

Table 6 and Table 7 present the results for the second hypothesis, which examines whether a strong CSR (proxied by HighCSR) moderates the market reaction to misconduct. Specifically, this model tests whether firms with higher ESG performance experience a less negative cumulative abnormal return following disclosure of tax violations, anti-money-laundering deficiencies, or accounting fraud.

In this specification for CAR_pm1, the interaction term treated:HighCSR is negative (-0.0031) but statistically insignificant ($p = 0.8158$). Similarly, for CAR_pm3 (Table 7), the interaction remains negative and again not statistically significant ($p = 0.1250$). This indicates that the presence of a strong CSR profile does not significantly buffer or reverse the effect of a misconduct announcement on short-term abnormal stock returns in this sample. Similarly, the main effect of HighCSR is also not significant ($p = 0.5970$), suggesting that a firm's CSR standing, by itself, does not drive a distinct change in returns in this context.

Consistent with the previous model, the control variable ROA remains positive and highly significant (estimate = 0.3244, $p < 0.001$), reaffirming that profitability is a key factor linked to short-term market performance. Other control variables, including firm size, leverage, and net income, show no significant relationship with abnormal returns in this model.

Taken together, these results provide no clear evidence that CSR performance alone softens the negative stock market impact of financial misconduct announcements, at least within the narrow event window and the sample used here. This underscores the importance of further exploring whether the severity of misconduct alters this relationship, which is tested in the next model.

5.6 Regression results of hypothesis 3

Table 8: Regression Results for H3 (CAR_pm1)

Variable	Estimate	Std. Error	t-Statistic	p-Value
(Intercept)	−0.0113	0.0391	−0.2899	0.7728
HighCSR	−0.1646	0.0620	−2.6549	0.0098
log_penalty_adjusted	−0.0009	0.0021	−0.4186	0.6768
Civil_Criminal_Numeric	−0.0105	0.0154	−0.6855	0.4953
log_at	−0.0001	0.0032	−0.0250	0.9801
roa	0.5396	0.0768	7.0262	0.0000
leverage	0.0139	0.0324	0.4303	0.6683
ni	0.0000	0.0000	0.6621	0.5101
HighCSR:log_penalty_adjusted	0.0104	0.0041	2.5346	0.0135
HighCSR:Civil_Criminal_Numeric	0.0235	0.0465	0.5049	0.6152

Table 9: Regression Results for H3 (CAR_pm3)

Variable	Estimate	Std. Error	t-Statistic	p-Value
(Intercept)	0.3423	0.0833	4.1081	0.0001
HighCSR	−0.2875	0.1323	−2.1739	0.0331
log_penalty_adjusted	−0.0065	0.0045	−1.4464	0.1525
Civil_Criminal_Numeric	0.0119	0.0328	0.3625	0.7181
log_at	−0.0204	0.0067	−3.0209	0.0035
roa	0.0714	0.1638	0.4355	0.6645
leverage	−0.0450	0.0691	−0.6510	0.5172
ni	0.0000	0.0000	1.7835	0.0788
HighCSR:log_penalty_adjusted	0.0160	0.0088	1.8331	0.0710
HighCSR:Civil_Criminal_Numeric	−0.0001	0.0992	−0.0013	0.9990

Table 8 and 9 present the regression results for the third hypothesis, which tests whether the moderating effect of CSR on the market reaction to financial misconduct depends on the perceived severity of the violation. Severity is captured both by the logarithm of the adjusted monetary penalty and by whether the case was handled as a civil or criminal matter.

For CAR_pm1 (table 8), the interaction between HighCSR and the size of the penalty (HighCSR:log_penalty_adjusted) is positive and statistically significant at the 5% level ($\beta = 0.0104$, $p = 0.0135$). This indicates that, contrary to the expectation that CSR loses its protective effect as the penalty increases, firms with higher CSR scores tend to show slightly less negative abnormal returns when the penalty amount is larger. However, the effect is small.

The interaction term for HighCSR:Civil_Criminal_Numeric, which captures whether the violation is criminal (1) or civil (0), is also positive but not statistically significant ($\beta = 0.0235$, $p = 0.61$). This suggests that there is no robust evidence that CSR moderates the difference in investor reaction between civil and criminal cases in this sample.

For CAR_pm3 (Table 9), a similar pattern is observed. The HighCSR:log_penalty_adjusted interaction is positive but only marginally significant ($p = 0.0710$), while the HighCSR:Civil_Criminal_Numeric interaction is clearly not significant.

Among the controls, firm profitability (ROA) remains strongly significant and positively related to abnormal returns ($\beta = 0.5396$, $p < 0.001$ for CAR_pm1). Other control variables do not show significant effects.

Overall, the results provide only partial support for H3. While the direction of the interaction with penalty size is significant, it does not align with the hypothesized negative relationship where CSR protection should weaken or backfire when misconduct is severe. The civil/criminal distinction does not appear to play a significant role as a conditional factor in this setting.

5.7 Summary of results

Overall, the results provide a nuanced but valuable insight into how financial markets respond to corporate misconduct and whether a strong CSR profile can moderate this reaction, especially when the violation is severe or criminal in nature.

For H1, the evidence does not indicate a consistent or significant direct impact of misconduct announcements on short-term stock returns. Across both the immediate (CAR_pm1) and slightly longer (CAR_pm3) windows, the treatment effect remains statistically insignificant. This suggests that, on average, investors do not systematically penalize firms for disclosures of tax violations, anti-money laundering deficiencies, or accounting fraud, at least not within the short event windows considered.

For H2, the findings offer no robust support for a protective effect of CSR. Although the coefficients for the interaction term between misconduct and HighCSR are negative for both CAR_pm1 and CAR_pm3, they are not statistically significant. This indicates that firms with stronger CSR scores do not appear to benefit from a meaningful reputational buffer when misconduct is revealed, challenging the idea of CSR functioning as a reliable form of reputational insurance in these contexts.

For H3, which tests whether the moderating effect of CSR weakens or reverses when violations are more severe or criminal, the results are mixed and partially counterintuitive. The interaction between CSR and penalty size shows a small but positive and occasionally significant effect, implying that higher penalties do not consistently erode the limited protective role of CSR, and may even slightly cushion market reactions. However, the interaction with the civil/criminal distinction is not significant, providing no evidence that legal severity alters the relationship meaningfully.

Taken together, these findings suggest that CSR's capacity to shield firms from market penalties in the event of financial misconduct is limited and highly dependent on its context. Moreover, the overall muted stock price response highlights that investors may already anticipate such events or perceive them as immaterial in light of broader firm fundamentals.

6 Discussion

This chapter reflects on the empirical results presented in Chapter 5 by connecting them back to the theoretical perspectives and prior research discussed earlier. The goal is to interpret what these findings imply about how financial markets react to announcements of financial misconduct, and whether a strong CSR profile can help mitigate these effects, particularly when the misconduct is severe or criminal in nature.

6.1 Market reaction to financial misconduct

The first hypothesis tested whether announcements of tax violations, anti-money laundering deficiencies or accounting fraud trigger negative abnormal stock returns in the short term. Contrary to expectations drawn from event studies like Davidson et al. (1994) and Kim et al. (2011), the results show no statistically significant market reaction in the immediate or slightly longer event windows.

From an investor behaviour perspective, this muted reaction could have several explanations. First, it may reflect information leakage or anticipation: markets might already factor in reputational risks if a firm's governance or sector is known to be prone to compliance issues. Second, the modest sample size and varying number of matched control firms likely reduce the statistical power to detect effects that might exist but are small in magnitude. Third, it is possible that modern investors weigh financial misconduct differently today, particularly when fines are modest relative to firm size or overshadowed by broader market trends.

Overall, the lack of strong evidence for H1 suggests that, at least within this sample and timeframe, short-term stock returns may be less sensitive to misconduct disclosures than prior studies imply, a finding that warrants caution when generalizing to other contexts or more severe scandals.

6.2 The moderating role of CSR

The second hypothesis explored whether firms with a stronger CSR profile are shielded from negative market reactions when misconduct becomes public. Here, the results do not show significant evidence of such a buffering effect. Across both event windows, the interaction term between CSR and misconduct is not significant.

This finding aligns with the mixed views in the literature. While Godfrey et al. (2009) and Minor and Morgan (2011) argue that CSR can act as "moral capital", more recent work (Bartov et al., 2019; Marques et al., 2025) cautions that stakeholders may feel betrayed when a self-declared "responsible" firm is caught behaving unethically. In this study, the absence of a clear protective effect suggests that the market may see CSR efforts as insufficient to counteract the reputational damage of clear compliance failures, or that CSR signals themselves may be discounted if they lack credibility.

Again, it is important to note that limited sample size and moderate variation in CSR scores might constrain the power to detect subtle moderation effects.

6.3 Conditional moderation: severity and criminality

The third hypothesis tested whether the (limited) protective effect of CSR weakens or reverses when misconduct is more severe (higher penalties) or when violations are criminal rather than civil. The results show no robust evidence that severity or criminality meaningfully alter the CSR–misconduct link. The three-way interaction terms for penalty size and the civil/criminal indicator are small and not consistently significant.

This outcome partly supports the argument that once misconduct crosses a reputational threshold, even a strong CSR profile cannot compensate for the perceived betrayal of trust (Marques et al., 2025). However, the insignificant findings may also stem from technical limitations: since control firms by design have no penalties, penalty-related variables apply only to the treated group, making precise estimation challenging in a small sample.

Despite this, the direction of some coefficients hints at a possible erosion of CSR's protective value when fines are higher, aligning with theoretical expectations that severe wrongdoing undermines moral capital.

6.4 Integrating theory and limitations

Taken together, these findings provide cautious but informative insights. From a signalling theory lens, CSR does not appear to reassure investors strongly when misconduct occurs. Signals may be seen as cheap talk if contradicted by concrete compliance failures. Agency theory partly explains the broader pattern: governance quality, not CSR labels alone, likely drives investors' trust in a firm's integrity. Stakeholder theory helps clarify why some firms invest in CSR to build reputational buffers, but this study suggests those buffers may be fragile when serious wrongdoing comes to light.

A key limitation is the dataset's modest power and sample size. As shown by the power checks, the analyses were underpowered to detect subtle effects, especially for interaction terms. Future research should replicate these tests with larger samples and richer data on the nature and timing of both CSR activities and misconduct events.

6.5 Practical implications

For managers and policymakers, the results underline that CSR alone is unlikely to fully shield firms from reputational or market fallout if compliance systems are weak. Firms should complement CSR initiatives with robust governance, transparent reporting, and credible enforcement of ethical standards. For investors, the findings highlight the need to critically evaluate CSR claims in light of actual conduct and governance practices.

While the study did not find strong statistical evidence supporting all hypotheses, it contributes to an ongoing conversation: CSR may help shape perceptions of a firm's values but cannot substitute for genuine compliance and sound governance.

7 Conclusion

This thesis investigated how financial markets respond to announcements of corporate financial misconduct, specifically tax violations, anti-money laundering deficiencies, and accounting fraud, and whether a strong Corporate Social Responsibility (CSR) profile can moderate these market reactions. Drawing on agency theory, signalling theory and stakeholder theory, the study developed and tested three hypotheses using a unique event-based dataset that combines violation records, stock market returns, and CSR performance indicators for publicly listed firms between 2010 and 2024.

The results provide limited evidence for the hypotheses. The first hypothesis, which predicted that firms experience negative abnormal returns following misconduct announcements, was not supported with statistical significance across the main event windows. Similarly, the second hypothesis that high CSR performance buffers negative market reactions did not receive clear support. Finally, the third hypothesis, which tested whether the protective effect of CSR weakens as misconduct severity or criminality increases, also yielded inconclusive results.

While these findings contrast with parts of prior literature that suggest clear reputational penalties and protective effects of CSR, they highlight the complex way in which investors process information about corporate misconduct in contemporary markets. The modest sample size, varying quality of control firm matching, and generally low statistical power are important constraints that may explain the absence of strong significance. Nonetheless, the observed patterns suggest theoretical relevance: severe penalties and criminal misconduct may weaken the credibility of CSR efforts, which aligns with expectations from signalling theory.

Overall, the study contributes to an ongoing discussion: CSR may shape how firms are perceived but cannot fully protect them from the reputational fallout of serious compliance failures. In practice, robust governance and ethical conduct appear more reliable than image-driven CSR in protecting investor trust.

7.1 Limitations and future research

As with any empirical study, several limitations should be acknowledged. First and foremost is the dataset's limited size and unbalanced matching structure: while every misconduct event was matched with at least one control firm, the number of controls varied, which reduces consistency in comparison. Additionally, the event study approach relies on clean, timely disclosures and assumes efficient market reactions, factors that can be blurred by media leaks or regulatory delays.

Another important limitation is statistical power. Despite careful matching and robust regression specifications, the relatively small sample and moderate variation in CSR scores limit the ability to detect subtle effects, especially for interaction terms involving severity and criminality.

Measurement also poses constraints. CSR is proxied by the ESG Combined Score, which captures broad reputational and governance signals but does not distinguish precisely between genuine governance practices and more symbolic disclosure. Similarly, the severity of misconduct is measured through monetary penalties, which may not fully reflect public perception of wrongdoing's seriousness.

Future research could build on this study by using larger multi-country samples, disaggregating CSR into its environmental, social, and governance subcomponents, and exploring how media tone and public sentiment interact with CSR credibility during misconduct events.

7.2 Practical implications

This thesis offers several takeaways for managers, investors, and policymakers. For companies, the results reinforce that CSR alone is not a foolproof shield against the reputational damage of unethical conduct. Firms should ensure that social responsibility initiatives are supported by rigorous internal governance and compliance systems.

For investors, the findings serve as a reminder to critically evaluate CSR claims alongside clear indicators of corporate integrity. High CSR scores may enhance reputation, but when paired with significant misconduct, they risk backfiring and eroding trust even further.

For regulators and standard setters, the study underscores the need for transparent, comparable disclosures not just of sustainability performance but also of governance practices and compliance histories. As sustainability and ESG reporting standards evolve, clearer links between social responsibility claims and robust corporate governance could help investors make better-informed decisions.

In sum, this thesis contributes to a nuanced understanding of how CSR interacts with financial misconduct in capital markets. While the protective role of CSR remains ambiguous under severe misconduct, the broader lesson is clear: authentic governance and ethical compliance remain central to safeguarding investor confidence in the long term.

8 References

Austin, C. R., & Wilson, R. J. (2017). An examination of reputational costs and tax avoidance: Evidence from firms with valuable consumer brands. Journal of the American Taxation Association, 39(1), 67-93.

Bartov, E., Marra, A., & Momenté, F. (2019). Corporate social responsibility and the market reaction to negative events: Evidence from inadvertent and fraudulent restatement announcements. The Accounting Review.

Beer, S., De Mooij, R. & Liu, L. (2020). International Corporate Tax Avoidance: A Review Of The Channels, Magnitudes, And Blind Spots. Journal of Economic Surveys, 34(3), 660 688.

Blaufus, K., Möhlmann, A., & Schwäbe, A. N. (2019). Stock price reactions to news about corporate tax avoidance and evasion. Journal of Economic Psychology, 72, 278-292.

Cooper, M., & Nguyen, N. (2020). Tax avoidance and stock price crash risk: Evidence from individual investors. International Business Review, 29, 101692.

Cooter, R. (1998). Expressive law and economics. Journal of Legal Studies, 27(S2), 585-607.

Davidson, W. N., Worrell, D. L., & Lee, C. I. (1994). Stock market reactions to announced corporate illegalities. Journal of Business Ethics, 13(13), 979-987.

Desai, M.A., Dharmapala, D. (2006). Corporate tax avoidance and high-powered incentives. Journal of Financial Economics, 79(1), 145-179

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. Journal of Financial Economics, 33(1), 3-56.

Godfrey, P. C., Merrill, C. B., & Hansen, J. M. (2009). The relationship between corporate social responsibility and shareholder value: An empirical analysis. Strategic Management Journal, 30(4), 425–445.

Graham, J. R., Hanlon, M., Shevlin, T., Shroff, N. (2014). Incentives for tax planning and avoidance. The Accounting Review, 89(3), 991-1023.

Hanlon, M., & Heitzman, S. (2010). A review of tax research. Journal of Accounting and Economics, 50(2-3), 127-178.

Hanlon, M., & Slemrod, J. (2009). What does tax aggressiveness signal? Evidence from stock price reactions to news about tax shelter involvement. Journal of Public Economics, 93(1-2), 126-141.

Hoi, C. K., Wu, Q., & Zhang, H. (2013). Is corporate social responsibility (CSR) associated with tax avoidance? Evidence from irresponsible CSR activities. *The Accounting Review, 88*(6), 2025-2059.

Ichev, R. (2023). Reported corporate misconducts: The impact on the financial markets. PLoS ONE, 18(2), e0276637.

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

Lanis, R., & Richardson, G. (2012). Corporate social responsibility and tax aggressiveness: An empirical analysis. Journal of Accounting and Public Policy, 31(1), 86-108.

Lins, K. V., Servaes, H., & Tamayo, A. (2017). Social capital, trust, and firm performance: The value of corporate social responsibility during the financial crisis. Journal of Finance, 72(4), 1785-1824.

Marques, M, Montenegro, T.M., & Brás, F.A. (2025). Tax Avoidance and Corporate Social Responsibility: A Meta-Analysis. Journal of American Taxation Association, 46(1), 137-156.

Minor, D., & Morgan, J. (2011). CSR as reputation insurance: Primum non nocere. California Management Review, 53(3), 40-59.

Mkadmi, A., & Ben Ali, S. (2024). How does tax avoidance affect corporate social responsibility and financial ratios in emerging economies. Journal of Economic Criminology, 5, 100070

Shiu, Y., & Yang, S. (2017). Does engagement in corporate social responsibility provide insurance like effects? Strategic Management Journal, 38(2), 455-470.

Sikka, P. (2010). Smoke and mirrors: Corporate social responsibility and tax avoidance. Accounting Forum, 34(3-4), 153-168.

Wang, F., Xu, S., & Sun, J. (2020). Corporate Tax Avoidance: A Literature Review And Research Agenda. Journal of Economic Surveys, 34(4), 793-811.