



School of Social Sciences

Master's Degree Programme in Business Administration

Postgraduate Dissertation

Material Disclosure Quality and Stock Return Volatility: Evidence
from Global Banking

Dimitrios Nektarios Alexandridis

Supervisor: Georgios Kolias

Patras, Greece, May 2026

Theses / Dissertations remain the intellectual property of students (“authors/creators”), but in the context of open access policy they grant to the HOU a non-exclusive license to use the right of reproduction, customisation, public lending, presentation to an audience and digital dissemination thereof internationally, in electronic form and by any means for teaching and research purposes, for no fee and throughout the duration of intellectual property rights. Free access to the full text for studying and reading does not in any way mean that the author/creator shall allocate his/her intellectual property rights, nor shall he/she allow the reproduction, republication, copy, storage, sale, commercial use, transmission, distribution, publication, execution, downloading, uploading, translating, modifying in any way, of any part or summary of the dissertation, without the explicit prior written consent of the author/creator. Creators retain all their moral and property rights.



Material Disclosure Quality and Stock Return Volatility: Evidence from Global Banking

Dimitrios Nektarios Alexandridis

Supervising Committee

Supervisor:

Georgios Kolias

University of Ioannina,
Hellenic Open University

Co-Supervisor:

Leonidas Doukakis

Athens University of Economics and
Business,
Hellenic Open University

Patras, Greece, May 2026

Abstract

This thesis investigates whether the quality of bank Environmental, Social and Governance (ESG) disclosures reduces stock return volatility in a global sample of 80 listed commercial banks drawn from the S&P Top 100 Banks ranking and spanning the period 2020-2024. Using computational text analysis on 390 bank-year reports (285 standalone ESG/sustainability documents and 105 integrated annual reports), four theoretically motivated disclosure-quality metrics are constructed: the Gunning Fog Index of readability; the Loughran-McDonald net positive sentiment ratio (tone); ESG vocabulary density; and a three-component SASB material-topic coverage metric capturing depth, balance and breadth across the six SASB Commercial Banking material topics. The dependent variable is annualised stock return volatility, computed from daily log returns. The regression specifications include year-varying market capitalisation, year-varying beta against MSCI ACWI, year-varying return on equity, the year-varying Tier 1 capital ratio, total assets, a report-type dummy, log word count, and year and region fixed effects, with standard errors clustered at the bank level. The headline findings strongly support the linguistic-credibility dimension of disclosure quality: the Loughran-McDonald tone of bank ESG reports is robustly and negatively associated with stock return volatility (beta = -0.122, $p = 0.001$ in the preferred cross-sectional specification, beta = -0.055, $p = 0.030$ once year-varying ROE and beta are added, and beta = -0.203, $p < 0.001$ in the standalone-reports subsample), is significant at the 1% level under Beck-Katz panel-corrected standard errors and feasible GLS with AR(1) errors alike, and survives explicit bank-size control and exclusion of the high-volatility COVID year. The Gunning Fog Index displays a positive association with volatility consistent with the obfuscation channel, significant at 5% in the cross-section (beta = +0.0072, $p = 0.022$) and in the standalone-reports subsample (beta = +0.0068, $p = 0.029$), and marginally significant at 10% in the full-control specifications and under all four alternative error-structure estimators. Both stylistic results are robust to adding the Tier 1 capital ratio as a solvency control, and to a battery of identification checks: a Mundlak correlated-random-effects decomposition with a regression-based Hausman test confirms that pooled OLS consistently estimates the disclosure effects, which load on the between-bank dimension (a test restricted to the disclosure variables does not reject pooled consistency, $F = 1.97$, $p = 0.107$); and under a one-year lag the readability effect is essentially unchanged — prior-year obfuscation predicts current-year volatility — whereas

the tone effect attenuates, indicating a contemporaneous association. The classical impression management hypothesis is partially supported: banks with lower ROE produce systematically less readable ESG reports ($\beta = -5.55$, $p = 0.028$), providing direct evidence of strategic obfuscation when financial performance is weak; the companion tone-inflation channel is not supported in either direction ($\beta = +0.366$, $p = 0.175$). The tone effect is, however, performance-contingent: a Tone x ROE interaction ($\beta = 0.717$, $p = 0.039$) shows that the volatility-dampening association of positive tone is concentrated among weak performers and is absent for strong performers, locating the tone channel precisely where impression-management incentives are strongest. A novel descriptive contribution emerges from a topic-conditioned analysis of disclosure quality: in 99.5% of bank-year reports the Fog Index is materially higher in sentences referencing SASB material topics than in the rest of the report (mean within-document gap of +9.5 Fog points, Cohen's $d = +1.76$), and in 91.3% of reports the Loughran-McDonald tone of material content is markedly less positive than that of non-material content (mean gap -0.27, Cohen's $d = -1.43$). This pattern of selective obfuscation and de-tonalisation of material content is documented for the first time in a banking context, is highly statistically significant under every test, and is consistent across years, report types and regions; the regression evidence indicates that it operates as a structural feature of the industry rather than a discriminator across banks. The substantive content metrics (ESG vocabulary density, SASB material-topic coverage) are not significantly associated with volatility in the cross-section, and the materiality premium hypothesis is not supported. The findings have implications for the design of mandatory sustainability disclosure frameworks (CSRD, ESRS, ISSB), suggesting that readability and tone standards should complement the substantive coverage requirements that dominate current regulatory practice, and that targeted constraints on material-topic readability may have particular regulatory leverage.

Keywords

ESG disclosure quality, stock return volatility, Loughran-McDonald sentiment, Gunning Fog Index, SASB materiality, signalling theory

“Γνωστοποίηση σημαντικών πληροφοριών, ποιότητα και μεταβλητότητα της απόδοσης των μετοχών: Στοιχεία από τον παγκόσμιο τραπεζικό τομέα”

“Δημήτριος Νεκτάριος Αλεξανδρίδης”

Περίληψη

Η παρούσα διατριβή διερευνά κατά πόσον η ποιότητα των γνωστοποιήσεων των τραπεζών σε θέματα Περιβάλλοντος, Κοινωνίας και Διακυβέρνησης (ESG) μειώνει τη μεταβλητότητα της απόδοσης των μετοχών, σε ένα παγκόσμιο δείγμα 80 εισηγμένων εμπορικών τραπεζών που προέρχονται από την κατάταξη S&P Top 100 Banks και καλύπτουν την περίοδο 2020-2024. Χρησιμοποιώντας υπολογιστική ανάλυση κειμένου σε 390 ετήσιες εκθέσεις τραπεζών (285 αυτόνομα έγγραφα ESG/βιωσιμότητας και 105 ολοκληρωμένες ετήσιες εκθέσεις), καταρτίζονται τέσσερις θεωρητικά τεκμηριωμένοι δείκτες ποιότητας γνωστοποίησης: ο δείκτης αναγνωσιμότητας Gunning Fog, ο δείκτης καθαρού θετικού συναισθήματος (τόνος) κατά Loughran-McDonald, η πυκνότητα λεξιλογίου ESG και ένας δείκτης κάλυψης ουσιωδών θεμάτων SASB τριών συνιστωσών που αποτυπώνει το βάθος, την ισορροπία και το εύρος στα έξι ουσιώδη θέματα εμπορικής τραπεζικής του SASB. Η εξαρτώμενη μεταβλητή είναι η ετησιοποιημένη μεταβλητότητα των αποδόσεων μετοχών, υπολογιζόμενη από τις ημερήσιες λογαριθμικές αποδόσεις. Οι προδιαγραφές της παλινδρόμησης περιλαμβάνουν την κεφαλαιοποίηση της αγοράς που ποικίλλει ανά έτος, το βήτα έναντι του MSCI ACWI που ποικίλλει ανά έτος, την απόδοση ιδίων κεφαλαίων που ποικίλλει ανά έτος, τον δείκτη κεφαλαίου κατηγορίας 1 (Tier 1) που ποικίλλει ανά έτος, τα συνολικά περιουσιακά στοιχεία, μια ψευδομεταβλητή τύπου εκθέσεως, τον λογαριθμικό αριθμό λέξεων, καθώς και σταθερές επιδράσεις έτους και περιοχής, με τυπικά σφάλματα ομαδοποιημένα (clustered) σε επίπεδο τράπεζας. Τα βασικά ευρήματα υποστηρίζουν σθεναρά τη διάσταση της γλωσσικής αξιοπιστίας όσον αφορά την ποιότητα της γνωστοποίησης: ο τόνος Loughran-McDonald των τραπεζικών εκθέσεων ESG συσχετίζεται σταθερά και αρνητικά με τη μεταβλητότητα της απόδοσης των μετοχών ($\beta =$

$-0,122$, $p = 0,001$ στην προτιμώμενη διατομεακή προδιαγραφή, $\beta = -0,055$, $p = 0,030$ όταν προστεθούν η ROE και το βήτα που διαφέρουν ανά έτος, και $\beta = -0,203$, $p < 0,001$ στο υποσύνολο των αυτόνομων εκθέσεων), είναι σημαντική στο επίπεδο του 1% τόσο υπό τα τυπικά σφάλματα διορθωμένα για δεδομένα πάνελ κατά Beck-Katz (PCSE) όσο και υπό τα εφικτά GLS (FGLS) με σφάλματα AR(1), και παραμένει σταθερή μετά από ρητό έλεγχο του μεγέθους της τράπεζας και τον αποκλεισμό του έτους της COVID με υψηλή μεταβλητότητα. Ο δείκτης Gunning Fog παρουσιάζει θετική συσχέτιση με τη μεταβλητότητα, σύμφωνη με το κανάλι ασάφειας, σημαντική στο επίπεδο του 5% στη διατομή ($\beta = +0,0072$, $p = 0,022$) και στο υποσύνολο των αυτόνομων εκθέσεων ($\beta = +0,0068$, $p = 0,029$), και οριακά σημαντική στο επίπεδο του 10% στις προδιαγραφές πλήρους ελέγχου και υπό όλους τους τέσσερις εναλλακτικούς εκτιμητές δομής σφάλματος. Και τα δύο υφολογικά ευρήματα είναι ανθεκτικά στην προσθήκη του δείκτη κεφαλαίου Tier 1 ως ελέγχου φερεγγυότητας, καθώς και σε μια σειρά ελέγχων ταυτοποίησης: μια αποσύνθεση συσχετισμένων τυχαίων επιδράσεων κατά Mundlak με παλινδρομικό έλεγχο Hausman επιβεβαιώνει ότι η συγκεντρωτική OLS (pooled OLS) εκτιμά με συνέπεια τις επιδράσεις της γνωστοποίησης, οι οποίες εδράζονται στη δια-τραπεζική (between-bank) διάσταση (ένας έλεγχος περιορισμένος στις μεταβλητές γνωστοποίησης δεν απορρίπτει τη συνέπεια της συγκεντρωτικής εκτίμησης, $F = 1,97$, $p = 0,107$)· και υπό χρονική υστέρηση ενός έτους, η επίδραση της αναγνωσιμότητας παραμένει ουσιαστικά αμετάβλητη — η ασάφεια του προηγούμενου έτους προβλέπει τη μεταβλητότητα του τρέχοντος έτους — ενώ η επίδραση του τόνου εξασθενεί, υποδηλώνοντας μια σύγχρονη (contemporaneous) συσχέτιση. Η κλασική υπόθεση διαχείρισης εντυπώσεων υποστηρίζεται εν μέρει: οι τράπεζες με χαμηλότερη απόδοση ιδίων κεφαλαίων (ROE) παράγουν συστηματικά λιγότερο ευανάγνωστες εκθέσεις ESG ($\beta = -5,55$, $p = 0,028$), παρέχοντας άμεση απόδειξη στρατηγικής ασάφειας όταν η χρηματοοικονομική απόδοση είναι αδύναμη· ο συνοδευτικός διάυλος διόγκωσης τόνου δεν υποστηρίζεται προς καμία κατεύθυνση ($\beta = +0,366$, $p = 0,175$). Ωστόσο, η επίδραση του τόνου εξαρτάται από την απόδοση: μια αλληλεπίδραση τόνου $\times$ ROE ($\beta = 0,717$, $p = 0,039$) δείχνει ότι η μειωτική για τη μεταβλητότητα συσχέτιση του θετικού τόνου συγκεντρώνεται στις τράπεζες με ασθενείς επιδόσεις και απουσιάζει για τις τράπεζες με ισχυρές επιδόσεις, εντοπίζοντας τον διάυλο του τόνου ακριβώς εκεί όπου τα κίνητρα διαχείρισης εντυπώσεων είναι ισχυρότερα. Μια νέα περιγραφική διαπίστωση προκύπτει από μια ανάλυση της ποιότητας της γνωστοποίησης με βάση τα θέματα: στο 99,5% των ετήσιων εκθέσεων των τραπεζών, ο δείκτης Fog είναι σημαντικά υψηλότερος

στις προτάσεις που αναφέρονται σε ουσιώδη θέματα του SASB σε σύγκριση με το υπόλοιπο της έκθεσης (μέση διαφορά εντός του εγγράφου +9,5 μονάδες Fog, d του Cohen = +1,76), ενώ στο 91,3% των εκθέσεων ο τόνος κατά Loughran-McDonald του ουσιώδους περιεχομένου είναι αισθητά λιγότερο θετικός από αυτόν του μη ουσιώδους περιεχομένου (μέση διαφορά -0,27, d του Cohen = -1,43). Αυτό το μοτίβο επιλεκτικής ασάφειας και αποτονικοποίησης του ουσιώδους περιεχομένου τεκμηριώνεται για πρώτη φορά σε τραπεζικό πλαίσιο, είναι στατιστικά εξαιρετικά σημαντικό σε κάθε δοκιμή και είναι συνεπές σε όλα τα έτη, τους τύπους εκθέσεων και τις περιοχές· τα στοιχεία της παλινδρόμησης υποδεικνύουν ότι λειτουργεί ως δομικό χαρακτηριστικό του κλάδου και όχι ως διακριτικό στοιχείο μεταξύ των τραπεζών. Οι δείκτες ουσιαστικού περιεχομένου (πυκνότητα λεξιλογίου ESG, κάλυψη ουσιωδών θεμάτων SASB) δεν συνδέονται σημαντικά με τη μεταβλητότητα στη διατομή, και η υπόθεση της πριμοδότησης της ουσιώδους πληροφόρησης δεν υποστηρίζεται. Τα ευρήματα έχουν επιπτώσεις στον σχεδιασμό των υποχρεωτικών πλαισίων γνωστοποίησης βιωσιμότητας (CSRD, ESRS, ISSB), υποδηλώνοντας ότι τα πρότυπα αναγνωσιμότητας και τόνου θα πρέπει να συμπληρώνουν τις απαιτήσεις ουσιαστικής κάλυψης που κυριαρχούν στην τρέχουσα ρυθμιστική πρακτική, και ότι οι στοχευμένοι περιορισμοί στην αναγνωσιμότητα των ουσιωδών θεμάτων ενδέχεται να έχουν ιδιαίτερη ρυθμιστική επίδραση.

Λέξεις – Κλειδιά

Ποιότητα γνωστοποίησης ESG, μεταβλητότητα αποδόσεων μετοχών, δείκτης Loughran-McDonald, δείκτης Gunning Fog, οικονομική σημαντικότητα SASB, θεωρία σηματοδότησης

Abstract	iv
Περίληψη.....	vi
List of Figures	xi
List of Tables.....	xii
List of Abbreviations & Acronyms	xiii
1. Introduction	1
1.1 Motivation	1
1.2 Research Gap	2
1.3 Research Objectives and Hypotheses.....	3
1.4 Main Findings	4
1.5 Contributions.....	7
1.6 Structure	8
2. Literature Review	10
2.1 ESG Disclosure and Capital Markets.....	10
2.2 Readability and Financial Markets.....	11
2.3 Textual Tone and Sentiment Analysis	12
2.4 SASB, Materiality, and Sector-Specific Disclosure	14
2.5 Disclosure in Banking: Regulatory and Empirical Context	15
3. Theoretical Framework and Hypotheses	17
3.1 Conceptual Framework and Hypotheses.....	17
3.1.1 Information Asymmetry and Volatility.....	17
3.1.2 Signalling Theory and Credibility.....	18
3.1.3 Obfuscation and Impression Management Hypotheses	18
3.2 Hypothesis Development	19
3.2.1 H1: Disclosure Quality and Stock Return Volatility.....	19
3.2.2 H2: Impression Management	20
3.2.3 H3: The Materiality Premium	21
3.3 Identification Strategy and Scope	22
4. Data & Methodology.....	23
4.1 Sample Construction	23
4.1.1 Universe and Inclusion Criteria	23
4.1.2 Report Classification.....	24
4.1.3 Sample Period	24
4.2 ESG Report Data Collection	24
4.3 Disclosure Quality Metrics	25
4.3.1 Fog Index	25
4.3.2 L-M Sentiment tone	25
4.3.3 ESG Vocabulary Density	26
4.3.4 SASB Material Topic Coverage.....	26
4.4 Financial and Market Data	27
4.4.1 Dependent Variable: Annualised Stock Return Volatility	27
4.4.2 Control Variables	27
4.4.3 Total Assets (Bank Size).....	29
4.4.4 Limitations of Financial Controls	29
4.5 Regression Models	30
4.5.1 Baseline Model (H1).....	30

4.5.2 Impression Management Model (H2)	31
4.5.3 Materiality Premium Model (H3)	31
4.5.4 Robustness Checks	31
5. Results	33
5.1 Descriptive Statistics	33
5.2 H1 – Disclosure Quality and Stock Return Volatility	37
5.3 H2 – Impression Management and Financial Performance	42
5.4 H3 – The Materiality Premium	45
5.5 Robustness	47
5.6 Summary of Empirical Results	56
6. Discussion & Conclusion	60
6.1 Summary of Findings and Theoretical Implications	60
6.2 Interpretation of the Tone Finding	62
6.3 Interpretation of the Asymmetric Impression Management Pattern	64
6.4 Limitations	65
6.5 Policy and Regulatory Implications	67
6.6 Future Research	68
6.7 Conclusion	70
References	72
Appendix A: “Sample Banks”	76
Appendix B: “SASB Material Topic Dictionary”	79
Appendix C: “Variable Definitions”	83
Appendix D: “Methodology Notes Register”	86

List of Figures

Figure 3.1: Conceptual Model

List of Tables

Table 1: Descriptive Statistics (Section 5.1)

Table 2: Correlation Matrix (Section 5.1)

Table 3: Main Regression Results (Section 5.2)

Table 4: Impression Management Regressions (Section 5.3)

Table 4B: Tone $\times$ Performance Interaction — Impression-Management Moderation
(Section 5.3)

Table 5: Materiality Premium Comparison (Section 5.4)

Table 6: Robustness Checks (Section 5.5)

Table 7: Alternative Error-Structure Estimators on Model 1c (Section 5.5)

Table 8: Selective Obfuscation, Fog and Tone in Material vs Non-Material Text (Section
5.5)

Table 9: Topic-Conditioned Disclosure Quality in the Volatility Model (Section 5.5)

Table 10: Mundlak Decomposition and Regression-Based Hausman Test (Section 5.5)

Table 11: One-Year Lag Robustness Table (Section 5.5)

List of Abbreviations & Acronyms

Abbreviation	Meaning
ACWI	All Country World Index (MSCI)
ADX	Abu Dhabi Securities Exchange
AR(1)	Autoregressive Process of Order 1
CSR	Corporate Social Responsibility
CSRD	Corporate Sustainability Reporting Directive
DXA	Document XML Attribute (unit, 1440 DXA = 1 inch)
ECB	European Central Bank
ESG	Environmental, Social, and Governance
ESRS	European Sustainability Reporting Standards
FE	Fixed Effects
FGLS	Feasible Generalised Least Squares
FX	Foreign Exchange
FY	Fiscal Year
GAAP	Generally Accepted Accounting Principles
HHI	Herfindahl-Hirschman Index
IFRS	International Financial Reporting Standards
ISSB	International Sustainability Standards Board
L-M	Loughran-McDonald
MSCI	Morgan Stanley Capital International
OCR	Optical Character Recognition
OLS	Ordinary Least Squares
PCSE	Panel-Corrected Standard Errors
PRB	Principles for Responsible Banking
ROE	Return on Equity
ROTE	Return on Tangible Equity
SASB	Sustainability Accounting Standards Board
SD	Standard Deviation
S&P	Standard & Poor's
TCFD	Task Force on Climate-related Financial Disclosures
URD	Universal Registration Document
USD	United States Dollar
VIF	Variance Inflation Factor

1. Introduction

1.1 Motivation

The past decade has seen a striking transformation in how large corporations, and banks in particular, communicate their non-financial performance. What began as voluntary social responsibility reporting in the early 2000s has evolved into a sophisticated industry: global banks now routinely publish ESG reports, CSR statements, and integrated annual reports running to hundreds of pages, covering everything from cyber security and diversity targets to governance structures and climate scenario analyses. The Global Reporting Initiative estimates that over 90% of large companies worldwide now produce some form of sustainability report. Regulatory mandates are accelerating this trend: the European Union's Corporate Sustainability Reporting Directive (CSRD), effective from 2024, requires detailed and audited sustainability disclosures from thousands of companies, while the International Sustainability Standards Board (ISSB) is driving convergence toward a global baseline. For banks, additional requirements, such as the Task Force on Climate-related Financial Disclosures (TCFD), the Principles for Responsible Banking (PRB), and Pillar 3 ESG risk requirements, have further broadened the disclosure ecosystem.

However, the explosion in sustainability reporting volume has not been matched by a corresponding consensus on quality. Sustainability report can range from detailed and informative accounts of a bank's material risks and stakeholder commitments to marketing exercises, replete with aspirational language, selective metrics, and strategically framed narratives. The same word count that signals transparency in one report signals obfuscation in another. The distinction between disclosure as genuine communication and disclosure as impression management lies at the heart of the present research.

This thesis asks: does the quality of a bank's ESG disclosure reduce stock return volatility? The motivation is grounded in market microstructure theory. Stock return volatility, at least in part, reflects the dispersion of investor beliefs about a firm's prospects and risk profile. When high-quality, credible information reaches the market, the range of plausible interpretations narrows and uncertainty declines. If ESG reports fulfil this function — communicating material risks clearly and credibly — we should observe measurably lower volatility for banks whose disclosures are readable, tonally measured, and substantively

complete. If, however, reports are deployed strategically to manage impressions rather than inform, the relationship may be attenuated or reversed.

Global banking provides an especially compelling setting for this inquiry. Banks are systemically important institutions whose stability has broad societal implications; investors, regulators, and civil society all have strong interests in the quality of bank disclosures. Banks also face a distinctive set of ESG material risks — financed emissions, data security, financial inclusion, and systemic risk management — that are not well captured by generic ESG frameworks. The Sustainability Accounting Standards Board (SASB) has identified six material ESG topics specifically for commercial banking, offering a sector-tailored benchmark against which actual report content can be assessed. Yet no large-scale study to date has tested whether banks' coverage of these material topics reduces investor uncertainty.

1.2 Research Gap

The academic literature on ESG disclosure and capital market outcomes is extensive and growing. A substantial body of evidence documents that companies with higher ESG scores or greater disclosure activity exhibit lower cost of capital, lower return volatility, and better long-run performance (Friede, Busse & Bassen, 2015; Lins, Servaes & Tamayo, 2017). However, this literature is predominantly focused on whether firms disclose rather than how well they do so. Several specific gaps motivate the present study:

First, disclosure quality is underspecified. Most empirical studies rely on third-party rating scores (Bloomberg ESG, MSCI, Refinitiv) or binary disclosure indicators as proxies for quality. These aggregated measures compress multiple dimensions — completeness, accessibility, tone, materiality — into a single score, obscuring which quality attributes investors actually price. A small but growing strand of research applies computational text analysis to financial disclosures (Li, 2008; Loughran & McDonald, 2011), but this work has focused almost entirely on annual reports and earnings communications rather than sustainability reports.

Second, the banking sector is underrepresented. The theoretical case for studying banks is strong: as financial intermediaries operating under heightened regulatory scrutiny, banks' disclosure quality has first-order implications for market discipline and financial stability. Yet empirical ESG studies routinely exclude financial firms or treat them as a single homogeneous sector, ignoring the substantial cross-bank heterogeneity in disclosure practice, regulatory environment, and business model.

Third, materiality has not been operationalised at the sector level. Khan, Serafeim & Yoon (2016) demonstrate that ESG performance on material issues — those identified as financially relevant to a sector — predicts future stock performance, while performance on immaterial issues does not. This materiality principle has not been extended to disclosure quality: we do not yet know whether disclosing material topics (not just performing on them) independently reduces investor uncertainty. The SASB Commercial Banking Standard makes such a test possible.

Fourth, the sample is global and the period is contemporary. Most banking disclosure studies are single-country or single-region. A panel spanning 80 banks across five global regions and five years (2020–2024) affords both cross-sectional richness and time-series variation encompassing the COVID-19 shock, the post-pandemic recovery, rising inflation, and the onset of mandatory CSRD reporting in Europe.

1.3 Research Objectives and Hypotheses

This thesis pursues three related lines of inquiry, formalised as testable hypotheses:

H1 — Disclosure Quality → Volatility: Higher-quality ESG disclosure is associated with lower annualised stock return volatility. Quality is operationalised through four text-analytic metrics: Fog Index (readability), L-M Tone (linguistic sentiment), ESG vocabulary density (breadth of sustainability content), and SASB material topic coverage (sector-specific completeness).

H2 — Impression Management: Banks with weaker financial performance produce more linguistically complex and more positively toned ESG reports, consistent with obfuscation

theory and the impression management hypothesis. Poorer performers are predicted to use higher Fog scores and more positive tone to mitigate negative market signals.

H3 — Materiality Premium: A regression model using SASB-specific material topic density explains a greater proportion of variance in stock return volatility than a model using generic ESG vocabulary density. This would confirm that the sector-specific materiality content of disclosure — not just its volume — is priced by investors.

1.4 Main Findings

The empirical analysis — based on pooled OLS regressions with year and region fixed effects and bank-level clustered standard errors, complemented by Beck-Katz panel-corrected standard errors and feasible GLS estimators, by a Mundlak correlated-random-effects decomposition with a regression-based Hausman test, and by one-year-lagged and interaction specifications as additional identification checks — yields six main findings. Throughout, the "full specification" denotes the augmented model that adds the Tier 1 capital ratio to the year-varying ROE and beta controls (Model 1b+T1), while the headline cross-sectional test remains the disclosure-quality-only specification (Model 1c).

L-M Tone is a robust cross-sectional predictor of volatility. A more positive net tone in a bank's ESG report is significantly associated with lower stock return volatility (beta = -0.122, $p = 0.001$ in the preferred cross-sectional specification on the 80-bank panel of 390 bank-year observations). The effect survives explicit bank-size control (beta = -0.130, $p = 0.001$), is amplified in dedicated standalone ESG reports (beta = -0.203, $p < 0.001$), survives exclusion of the COVID-2020 year (beta = -0.098, $p = 0.012$), and is also significant in the Asia-Pacific subsample. It remains significant at the 5% level when year-varying beta and ROE are jointly included (beta = -0.055, $p = 0.030$) and when the Tier 1 capital ratio is added in the augmented specification (beta = -0.057, $p = 0.024$), and it is significant at the 1% level under all four estimation strategies examined. A Mundlak decomposition shows that this effect operates on the between-bank dimension, and a regression-based Hausman test restricted to the disclosure variables does not reject the consistency of pooled

OLS ($F = 1.97$, $p = 0.107$); under a one-year lag, however, the tone effect attenuates and loses significance, indicating that the tone-volatility association is contemporaneous rather than predictive.

Fog Index displays a positive association with volatility consistent with the obfuscation channel. The effect is significant at the 5% level in the cross-sectional specification ($\beta = +0.0072$, $p = 0.022$) and in the standalone-reports subsample ($\beta = +0.0068$, $p = 0.029$), and is marginally significant at the 10% level under the full-control specifications including year-varying β and ROE ($\beta = +0.0034$, $p = 0.078$ in Model 1b) and after adding the Tier 1 capital ratio ($\beta = +0.0033$, $p = 0.085$ in Model 1b+T1). Under the four alternative error-structure estimators it is significant at the 1% level under Beck-Katz PCSE and FGLS with panel heteroskedasticity, at the 5% level under OLS with clustered standard errors, and at the 10% threshold under FGLS with AR(1) errors. Critically, the readability effect is essentially unchanged when the report-derived variables are lagged one year ($\beta = +0.0044$ versus $+0.0045$ contemporaneous): because current volatility cannot have caused a report written a year earlier, this provides forward-looking, reverse-causality-resistant evidence for the readability channel — a property the tone channel does not share.

The impression management hypothesis is partially supported through the readability channel and is not supported through the tone channel. Banks with lower return on equity produce systematically less readable ESG reports ($\beta = -5.55$, $p = 0.028$), providing direct evidence of strategic obfuscation when financial performance is weak. The companion tone-inflation channel is not supported: the coefficient on ROE in the tone regression is positive but insignificant ($\beta = +0.366$, $p = 0.175$). The two stylistic channels of impression management thus operate independently — readability tracks performance, tone does not — and the strong tone-volatility relationship documented for H1b should not be interpreted as the result of a performance-driven tone-inflation strategy.

The tone effect is performance-contingent. Although weak performers do not on average adopt more positive tone, the volatility-dampening association of positive

tone is concentrated among them: a Tone x ROE interaction is positive and significant ($\beta = +0.717$, $p = 0.039$), and the marginal effect of tone on volatility is large and significant for weak performers (-0.086 at one standard deviation below mean ROE, $p = 0.007$) but statistically indistinguishable from zero for strong performers (-0.023 , $p = 0.353$). Report tone therefore shapes investor uncertainty precisely where it is greatest — among weaker performers — and is uninformative where fundamentals are self-evident, locating the tone channel where impression-management incentives are strongest and complementing the direct H2 evidence.

SASB material topic coverage does not independently predict volatility in the cross-section, and the materiality premium hypothesis (H3) is not supported. The high collinearity between ESG vocabulary density and SASB density ($r = 0.854$) leaves the two density measures statistically indistinguishable in their effect on volatility. A within-bank result of secondary interest emerges in the Mundlak decomposition (SASB Density within-bank coefficient $+0.0165$, $p < 0.01$; corroborated under standard bank fixed effects at $+0.029$ and random effects at $+0.018$), most plausibly reflecting reverse causality whereby banks expand material-topic disclosure in more turbulent years. A topic-conditioned analysis described in Section 5.5 confirms that the cross-sectional null on the SASB metrics reflects a deeper structural pattern in the data rather than a measurement artefact.

A novel descriptive finding emerges from a within-document analysis of disclosure quality. In 388 of the 390 reports in the sample (99.5%), the Fog Index of sentences referencing SASB material topics is higher than the Fog Index of the rest of the report; the average gap is $+9.5$ Fog points, with a Cohen's d of $+1.76$ corresponding to a very large effect size. In 356 of 390 reports (91.3%), the Loughran-McDonald tone of material content is less positive than the tone of non-material content; the average gap is -0.27 , with a Cohen's d of -1.43 . Banks therefore systematically write more obscurely and in more measured tone about their material risks than about the remainder of their ESG narratives. The pattern is present in every year of the sample and in both standalone and integrated reports. It is, however, a level effect that operates near-universally across the industry rather than a differentiator across banks: cross-bank variation in the within-document gap does not independently

predict stock return volatility once the whole-document Fog and Tone measures are included.

These findings are broadly consistent with information asymmetry theory: the linguistic credibility of disclosure (tone) and its accessibility (readability) jointly shape investor uncertainty, while substantive content as measured by topic coverage does not provide additional explanatory power in the cross-section.

1.5 Contributions

This thesis makes five contributions to the literature.

Multi-dimensional quality measurement. Rather than relying on a single proxy, this study constructs four theoretically motivated text-analytic metrics from the raw content of 390 bank-year ESG reports (plus additional reports for unlisted banks processed through the same pipeline but absent from the regression panel), enabling a disaggregated assessment of the quality dimensions that matter for capital-market outcomes.

SASB operationalisation. The thesis introduces a redesigned three-component SASB coverage metric — combining keyword density, topical balance (HHI-based), and proportional coverage — that goes beyond simple dictionary hit counts and is suitable for cross-sectional regression analysis.

Global banking panel. A sample of 80 listed banks across five regions and five years, with year and region fixed effects, represents one of the largest ESG text-analysis datasets assembled for the banking sector. The 80 banks are drawn from the S&P Dec 2024 Top 100 ranking, making the sample broadly representative of the global commercial-banking industry.

Impression management test. The study provides direct evidence on whether bank managers use ESG reports strategically in response to financial performance. The asymmetric pattern — obfuscation operates as predicted, tone inflation does not — is itself a contribution and has direct relevance for the credibility and regulatory design of sustainability disclosure. The thesis further shows that the tone effect is performance-contingent: although weak performers do not on average inflate their tone, the volatility-dampening association of positive tone is concentrated among them (Tone x ROE interaction, $\beta = +0.717$, $p = 0.039$), locating the tone channel where impression-management incentives are strongest. The identification of these effects is supported by a Mundlak/Hausman analysis that justifies the pooled estimator and by one-year-lag specifications that distinguish the forward-looking readability channel from the contemporaneous tone channel.

Documentation of selective obfuscation. The topic-conditioned analysis of disclosure quality reported in Sections 5.1 and 5.5 documents, for the first time in a global banking sample, a systematic and very large within-document obfuscation and de-tonalisation of material content. In 99.5% of reports the Fog Index of material sentences exceeds the Fog Index of non-material sentences by an average of 9.5 Fog points; in 91.3% of reports the Loughran-McDonald tone of material content is less positive than the tone of non-material content by an average of 0.27 units. Both gaps are highly statistically significant and exhibit very large effect sizes (Cohen's $d > 1.4$). This finding has direct implications for sustainability disclosure regulation, suggesting that quality standards targeted specifically at material-topic readability may have greater regulatory leverage than uniform whole-document readability requirements.

1.6 Structure

Chapter 2 reviews the relevant literature across five bodies of work: ESG disclosure and capital markets, readability and financial communication, textual tone and sentiment analysis, SASB and materiality, and banking-specific disclosure. Chapter 3 develops the

theoretical framework and formally derives the three hypotheses. Chapter 4 describes the sample construction, data sources, metric computation, and estimation strategy in detail. Chapter 5 presents and interprets the empirical results. Chapter 6 discusses the findings, acknowledges limitations, and identifies avenues for future research.

2. Literature Review

2.1 ESG Disclosure and Capital Markets

The relationship between corporate disclosure and capital market outcomes has been a central concern in financial economics for decades. The foundational insight — that information asymmetry between managers and outside investors creates costs — was formalised by Akerlof (1970) in the context of adverse selection, and extended to voluntary disclosure by Grossman (1981) and Milgrom (1981), who showed that in equilibrium, firms with favourable private information have incentives to disclose it. Spence (1973) provided a broader signalling framework: voluntary actions, including disclosure, can credibly convey quality when the cost of the signal is inversely related to the signaller's quality. Applied to sustainability reporting, this framework predicts that high-quality, credible ESG disclosures should reduce the information gap between banks and their investors, lowering the uncertainty premia embedded in stock prices.

The empirical literature broadly supports a link between ESG disclosure and reduced market uncertainty. Diamond and Verrecchia (1991) demonstrate theoretically that voluntary disclosure reduces information asymmetry and improves market liquidity by attracting informed trading. Subsequent empirical work has found that firms with greater voluntary disclosure exhibit lower bid-ask spreads, lower cost of equity, and lower return volatility — all consistent with reduced adverse selection (Leuz & Verrecchia, 2000; Healy & Palepu, 2001). In the ESG context, Friede, Busse & Bassen (2015) synthesise over 2,000 empirical studies and report that approximately 90% find a non-negative relationship between ESG measures and financial performance, with the majority finding a positive association. More recently, Lins, Servaes & Tamayo (2017) document that high-CSR firms experienced significantly better stock returns during the 2008–2009 financial crisis, attributing this to the role of social capital in sustaining investor trust under conditions of extreme uncertainty.

However, this predominantly positive picture is complicated by a growing awareness that ESG disclosure is not uniformly informative. Clarkson, Li, Richardson & Vasvari (2008) show that voluntary environmental disclosures are driven partly by legitimacy-seeking motives: firms with poor environmental performance disclose strategically to manage external perceptions, creating a wedge between disclosure and underlying performance.

Lyon and Maxwell (2011) review the greenwashing literature, documenting systematic patterns of selective emphasis, misleading framing, and omission of adverse information in corporate environmental communications. In the banking context, the concern is particularly acute: banks can choose which aspects of their financed emissions, lending practices, or governance structures to highlight, with limited mandatory verification. The implication is that the market-dampening effect of ESG disclosure depends critically on whether that disclosure is genuinely informative or strategically managed — a distinction that motivates the quality-focused approach of the present study.

A key unresolved issue in the literature concerns the distinction between disclosure quantity and quality. Studies using binary disclosure indicators or rating scores may capture the decision to disclose rather than the value of what is disclosed. Marquis and Toffel (2014) demonstrate that firms strategically emphasise positive aspects of their performance while downplaying negatives, producing selective disclosures that do not fully reduce information asymmetry. This suggests that simple disclosure scores — which reward breadth of coverage but not depth or credibility — may be poor proxies for the quality attributes that investors actually price.

2.2 Readability and Financial Markets

The linguistic accessibility of corporate communications has attracted growing attention in accounting and finance. Gunning (1952) developed the Fog Index — a syllable-based measure of text complexity — in the context of business writing. Li (2008) introduced this measure to accounting research, demonstrating that the readability of annual reports (measured by Fog Index) is negatively related to earnings quality and positively related to the likelihood of earnings persistence: firms with more readable 10-Ks tend to have higher future earnings, suggesting that opaque language may signal managerial uncertainty or strategic obfuscation. Lehavy, Li & Merkley (2011) extend this finding to analyst behaviour, showing that firms with less readable annual reports attract greater analyst following — presumably because complex disclosures require more processing by information intermediaries.

The obfuscation hypothesis — that managers deliberately increase textual complexity when performance is poor — has received consistent empirical support. Courtis (1998) documents negative correlations between firm profitability and the readability of annual report narratives. Lo, Ramos & Rogo (2017) find that managers increase the linguistic complexity of earnings communications specifically in periods of unexpected earnings declines, consistent with deliberate obscuration of bad news. Bloomfield (2002) develops an "incomplete revelation hypothesis" suggesting that markets do not fully impound information that is costly to process, implying that investors may be less attentive to hard-to-read disclosures and thus exhibit more uncertainty about the underlying fundamentals.

These findings have motivated a growing literature on the readability of sustainability reports specifically. Melloni, Caglio & Perego (2017) analyse the readability of integrated reports and find that firms with poorer performance produce less readable integrated reports, consistent with the obfuscation hypothesis transposed to the ESG context. Caglio, Melloni & Perego (2020) further show that the readability of integrated reports is positively associated with reporting quality as assessed by the International Integrated Reporting Council's framework. For the banking sector, where ESG reports are largely voluntary and subject to little independent verification, the scope for strategic readability management may be considerable. If investors perceive opaque ESG reports as concealing adverse information, less readable disclosures should be associated with greater investor uncertainty and thus higher volatility — the prediction embedded in H1.

Despite the rich annual-report literature, no published study to date has directly examined the association between ESG report readability (rather than annual report readability) and stock return volatility for a global banking sample. The unique characteristics of sustainability reports — their voluntary nature, their explicit sustainability framing, their lack of standardised structure until recently — make this an open empirical question.

2.3 Textual Tone and Sentiment Analysis

Parallel to the readability literature, a substantial body of work has developed computational tools for analysing the linguistic tone of financial documents. Early sentiment analysis in finance relied on general-purpose dictionaries such as the Harvard General Inquirer, but

Loughran and McDonald (2011) demonstrated that word lists designed for general English are poorly suited to financial contexts (see also El-Haj, Rayson, Walker, Young & Simaki, 2019, for a survey of subsequent developments in computational analysis of financial discourse): common financial terms such as "liability," "tax," "loss," and "risk" are classified as negative by general dictionaries but carry neutral or technical meanings in financial discourse. Their purpose-built financial sentiment dictionary — hereafter the L-M dictionary — has become the standard tool in accounting and finance text analysis, with positive and negative word lists validated against financial document corpora; Loughran and McDonald (2016) review the broader textual-analysis literature this work has spawned.

The L-M dictionary has been applied across a wide range of financial communications. Loughran and McDonald (2011) show that their tone measure is significantly related to annual report filing date returns and subsequent stock volatility. Price, Doran, Peterson and Bliss (2012) apply sentiment analysis to earnings conference call transcripts and find that managerial tone predicts earnings surprises beyond analyst consensus. Huang, Teoh & Zheng (2014) document that optimistic tone in MD&A sections predicts future stock returns, with more optimistic-toned disclosures followed by weaker subsequent performance — consistent with strategic over-optimism when performance is about to deteriorate.

In the ESG context, the tone question takes on particular significance. ESG reports are structurally oriented toward positive framing — they celebrate achievements, announce targets, and position the firm as a responsible corporate citizen. The mean L-M Tone in our sample is 0.317, indicating that positive words outnumber negative words by approximately three to one. The question is whether this positivity is uniformly uninformative, or whether cross-sectional variation in tone conveys genuine signals about underlying ESG performance quality. Drempetic, Klein & Zwergel (2020) examine ESG report language and find that firms with stronger ESG performance use more specific and less hedged language — tentatively suggesting that tone variation is at least partly informative about underlying quality.

For the present study, the key theoretical tension is whether more positive tone in bank ESG reports is associated with lower volatility (credible signalling of good ESG performance) or higher volatility (credibility loss from uniformly positive framing that investors discount).

The direction of this effect is ultimately an empirical question; the sign of the tone coefficient in the regression is interpretively important and is discussed in Chapter 6.

2.4 SASB, Materiality, and Sector-Specific Disclosure

A critical development in the sustainability accounting literature has been the emergence of industry-specific materiality standards. Departing from earlier frameworks that prescribed uniform disclosure across all sectors, the Sustainability Accounting Standards Board (SASB) developed industry-specific standards that identify which ESG topics are likely to be financially material for each sector — based on evidence of their relevance to financial condition and operating performance. The Commercial Banking standard identifies six material topics (Sustainability Accounting Standards Board, 2023): (1) Data Security, (2) Financial Inclusion and Capacity Building, (3) Incorporation of ESG Factors in Credit Analysis, (4) Financed Emissions and Climate Risk, (5) Business Ethics, and (6) Systemic Risk Management.

The empirical significance of this materiality approach was demonstrated by Khan, Serafeim & Yoon (2016) in their landmark study. Constructing firm-level measures of performance on SASB-defined material versus immaterial ESG issues, they find that material ESG performance is positively and significantly associated with future abnormal stock returns, while immaterial ESG performance has no such association. This result has important implications for disclosure quality research: if investors price material ESG performance, they may similarly attend more closely to material ESG disclosure, treating it as a more credible and relevant signal than broad-based ESG language. This is the materiality premium hypothesis (H3) in the present study.

To our knowledge, no prior study has tested whether the coverage of SASB material topics in sustainability reports — as distinct from performance on those topics — differentially predicts stock return volatility relative to general ESG vocabulary coverage. The challenge is partly methodological: existing approaches typically measure disclosure quality through third-party ratings or binary topic coverage checklists, neither of which captures the density and balance of material topic discussion within report text. This thesis addresses the gap by

constructing a three-component SASB metric from raw report text, as described in Chapter 4.

It is worth noting that the six SASB material topics for commercial banking reflect the sector's specific systemic importance. Financed emissions recognise that banks' primary climate exposure is indirect — through the carbon profile of their lending books — rather than through their own operations. Systemic risk management reflects regulators' and investors' concern about banks' contributions to financial system fragility. These topics are not adequately captured by generic ESG frameworks, reinforcing the case for a sector-specific analysis.

2.5 Disclosure in Banking: Regulatory and Empirical Context

Banking is a distinctive setting for disclosure research because information asymmetry is exceptionally severe, the consequences of disclosure failures are potentially systemic, and the regulatory framework is unusually dense. Flannery and Thakor (2006) review the theoretical and empirical literature on bank transparency, emphasising that market discipline — the mechanism through which creditors and shareholders monitor banks — requires credible and timely information. Pillar 3 of the Basel III framework mandates quantitative disclosures of risk exposures precisely to enable this discipline; the implicit assumption is that disclosure reduces uncertainty and constrains risk-taking through market pressure.

The empirical literature on bank transparency and stock market outcomes provides the closest precedents for the present study. Nier and Baumann (2006) document that higher Pillar 3 disclosure quality is associated with lower risk-taking among international banks, consistent with disclosure reducing moral hazard. Morgan (2002) shows that credit rating agencies disagree more about bank bond ratings than about ratings for non-financial firms, reflecting the intrinsic opacity of bank balance sheets. This suggests that incremental information from voluntary disclosures — including ESG reports — may be especially valuable in the banking context, where mandatory financial disclosures leave substantial residual uncertainty.

The COVID-19 pandemic provides an important temporal context for the present sample. Duan, Li & Xue (2022) and related studies document sharp increases in bank stock return volatility in 2020 associated with uncertainty about loan book quality, economic disruption, and liquidity conditions. Year fixed effects in our regression models absorb this common shock, but the 2020 observations represent a period in which the informational value of disclosure may have been particularly salient — investors faced exceptional uncertainty, and any credible communication about a bank's resilience and risk management practices could plausibly have been priced.

The CSRD and ISSB developments in 2023–2024 also deserve attention as a structural shift in the reporting landscape. The CSRD mandates assured sustainability disclosure for large EU companies under the European Sustainability Reporting Standards (ESRS), moving ESG communication from voluntary to regulated territory for European banks in the sample. Year fixed effects partially absorb this shock, but the 2023–2024 observations for European banks involve a qualitatively different reporting regime from earlier years — a heterogeneity we acknowledge in the robustness discussion.

3. Theoretical Framework and Hypotheses

3.1 Conceptual Framework and Hypotheses

The theoretical architecture of this thesis rests on three interconnected foundations: information asymmetry theory, signalling theory, and the obfuscation (or impression management) hypothesis. Together, these frameworks generate predictions about both the market consequences of ESG disclosure quality (H1 and H3) and the strategic incentives that shape it (H2).

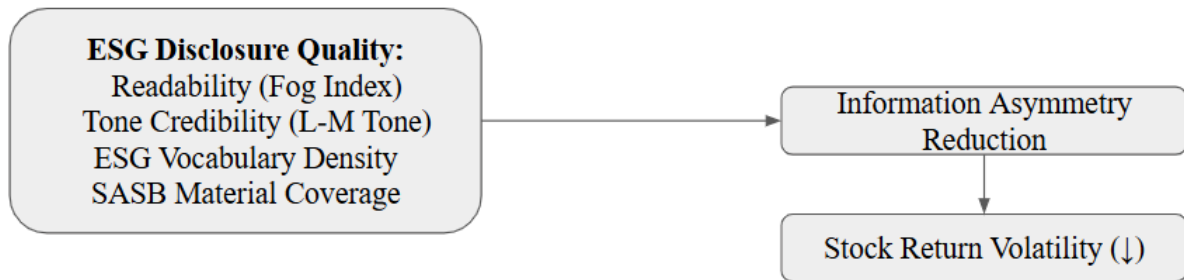
3.1.1 Information Asymmetry and Volatility

In financial markets, stock return volatility is a function of two distinct sources of uncertainty: genuine fundamental uncertainty about a firm's future cash flows, and information-based uncertainty arising from heterogeneous investor beliefs about current fundamentals. Whereas the former cannot be reduced by disclosure, the latter can. The Diamond-Verrecchia (1991) model formalises this distinction: voluntary disclosure reduces the adverse selection component of the bid-ask spread by narrowing the information advantage of informed traders, improving liquidity and reducing the return volatility attributable to information asymmetry.

Applied to sustainability reporting, this framework predicts that ESG disclosure quality should matter independently of the underlying quality of a bank's ESG performance. Even if two banks have identical underlying sustainability profiles, the bank whose ESG report is more readable, more tonally measured, and more complete in its coverage of material topics will leave investors with less residual uncertainty. Put differently, high-quality disclosure converts latent information into actionable knowledge more efficiently, compressing the dispersion of investor beliefs and thus dampening return volatility.

This channel implies a specific relationship: the quality attributes of disclosure — not merely the fact of disclosure — should predict volatility. A bank that discloses prolifically but obscurely (high Fog, generic ESG vocabulary, surface-level SASB coverage) provides less genuine information reduction than one that discloses concisely, clearly, and on the topics that matter most to investors. Figure 3.1 below summarises the proposed causal chain.

Figure 3.1: Conceptual Model



3.1.2 Signalling Theory and Credibility

Spence's (1973) signalling framework introduces an important qualification: for a disclosure signal to be credible, it must be costly to imitate. If high-quality disclosure — readable, materially complete, tonally measured — can be produced by any bank regardless of its underlying ESG performance, it loses its informational value. Several mechanisms may preserve credibility in practice. Readable, specific disclosures are easier for analysts and civil society organisations to scrutinise and challenge; banks making specific, quantified commitments face reputational costs if they fall short. The tone finding in our data — that more positive tone is associated with lower volatility — may reflect either genuine signalling (banks with good performance are confident enough to use positive language) or a market convention in which investors attend to the presence of negative or cautious language as the primary signal of underlying problems.

3.1.3 Obfuscation and Impression Management Hypotheses

A complementary literature in management accounting proposes that disclosure style is not merely a passive reflection of underlying performance but an active managerial choice. The obfuscation hypothesis (Courtis, 1998; Lo et al., 2017) posits that managers strategically deploy linguistic complexity to reduce the accessibility of adverse information: when performance is poor, opaque text slows the processing of bad news. The impression management hypothesis (Clatworthy & Jones, 2003; Brennan, Guillamon-Saorin & Pierce, 2009) extends this: managers use the full stylistic repertoire — tone, emphasis, narrative sequence, graphic design — to shape the impressions formed by report readers.

In the ESG context, impression management takes on particular relevance because sustainability reports are overwhelmingly voluntary, narratively structured, and relatively free of mandatory quantitative constraints. Unlike financial statements, which are anchored to audited numbers, ESG reports offer wide latitude for narrative framing. A bank facing regulatory scrutiny over its lending practices may emphasise its financial inclusion initiatives; one with poor climate performance may foreground governance improvements. This strategic dimension implies that the relationship between financial performance and disclosure quality may itself be endogenous: disclosure style is partly a function of what the bank wishes to communicate, not only of its underlying performance.

3.2 Hypothesis Development

3.2.1 H1: Disclosure Quality and Stock Return Volatility

The information asymmetry framework implies that higher-quality ESG disclosure reduces the component of stock return volatility attributable to investor uncertainty. Each of the four quality metrics captures a distinct mechanism through which this reduction operates.

Readability. A less readable report — higher Fog Index — increases the processing cost of information extraction for investors. Under Bloomfield's (2002) incomplete revelation hypothesis, information that is difficult to process is less fully incorporated into prices, leaving greater residual uncertainty. Furthermore, less readable disclosures may signal managerial uncertainty or strategic obfuscation, compounding investor uncertainty. We therefore predict:

H1a: The Fog Index is positively associated with stock return volatility (less readable reports → higher volatility).

Linguistic Tone. The relationship between tone and volatility is theoretically ambiguous. On one hand, sustained positive sentiment in ESG reports may reflect genuine good performance and credible optimism, narrowing the distribution of investor assessments. On the other hand, uniformly high positive tone may be discounted as boosterism, increasing rather than decreasing uncertainty. The net empirical prediction depends on whether the signalling or credibility-loss channel dominates. Given the evidence in Loughran and

McDonald (2011) that tone is informative in financial texts, we hypothesise a negative association:

H1b: L-M Tone is negatively associated with stock return volatility (more positive tone → lower volatility), conditional on the credibility of the signal.

ESG Vocabulary Density and SASB Coverage. Greater ESG content density — more sustainability-relevant language relative to total report length — signals substantive engagement with non-financial topics. SASB material topic coverage goes further by capturing the depth of discussion on the specific topics identified as financially relevant for commercial banking. Both metrics should reduce investor uncertainty about a bank's non-financial risk profile, though we expect SASB coverage to have a stronger effect if investors attend preferentially to material topics:

H1c: ESG vocabulary density and SASB material topic coverage are negatively associated with stock return volatility.

Taken together:

H1 (overall): Higher ESG disclosure quality — operationalised as lower Fog Index, more positive L-M Tone, greater ESG vocabulary density, and greater SASB material topic coverage — is associated with lower annualised stock return volatility, controlling for bank size, financial performance, systematic risk, report type, and time and region fixed effects.

3.2.2 H2: Impression Management

The impression management hypothesis predicts that financially weaker banks will strategically alter the style of their ESG reports to manage external perceptions. Two mechanisms are theorised:

Obfuscation through readability. When return on equity is low, banks may increase report complexity to slow the processing of adverse performance signals, reducing the probability that sophisticated readers reach conclusions unfavourable to management. This implies a negative relationship between ROE and Fog Index: lower financial performance → higher Fog.

Positive tone inflation. Alternatively, weaker performers may rely on aspirational language and positively framed narratives to offset negative signals from financial results. This implies a negative relationship between ROE and L-M Tone: lower ROE → more positive tone. Note that this mechanism runs counter to the signalling prediction in H1b — if positive tone in poorer-performing banks reflects strategic inflation rather than genuine confidence, the tone signal is contaminated, potentially explaining why the overall tone effect on volatility is not stronger.

H2: Banks with lower return on equity produce ESG reports with higher Fog Index scores and more positive L-M Tone, consistent with linguistic obfuscation and strategic positive framing when financial performance is weak.

3.2.3 H3: The Materiality Premium

Khan, Serafeim & Yoon (2016) establish that performance on SASB material issues is priced by markets while performance on immaterial issues is not. By analogy, if investors attend selectively to material ESG disclosures, then a model of stock return volatility including SASB material topic coverage should fit the data better than one relying solely on generic ESG vocabulary density — the materiality premium hypothesis.

Formally, let Model A specify volatility as a function of ESG vocabulary density (plus controls) and Model B specify it as a function of SASB density (plus controls). H3 predicts that Model B achieves higher R^2 than Model A, and that the standardised coefficient on SASB density exceeds that on ESG density in Model A, reflecting investors' preferential attention to sector-relevant material content.

H3: ESG report coverage of SASB-defined material commercial banking topics explains a greater proportion of variance in stock return volatility than generic ESG vocabulary density, reflecting a materiality premium in the informational value of sustainability disclosure.

3.3 Identification Strategy and Scope

Several identification challenges are acknowledged. First, the relationship between disclosure quality and volatility may be partly reverse causal: banks with lower fundamental volatility may be more confident in the quality of their disclosures, or may have fewer adverse events to hide. This threat is partially mitigated by the inclusion of beta (a forward-looking measure of systematic risk) and market capitalisation as controls, and by the robustness check excluding the COVID year 2020 when volatility was predominantly exogenously determined. Full causal identification would require an instrument or a natural experiment — an avenue for future work.

Second, the panel structure raises concerns about serial correlation in bank-level residuals. All models cluster standard errors at the bank level and include year fixed effects to absorb common time shocks. Region fixed effects account for systematic regulatory and cultural differences in disclosure practices. The combination of cross-sectional richness (80 banks, five regions) and temporal variation (five years, including a structural break in 2020) provides reasonable identification power for cross-sectional quality effects.

4. Data & Methodology

4.1 Sample Construction

4.1.1 Universe and Inclusion Criteria

The sample is drawn from the S&P Global Top 100 Banks by Total Assets ranking (December 2024 edition), filtered to the universe of banks for which usable English-language sustainability reports are available across the 2020–2024 window. The initial universe of 100 banks is filtered as follows.

Excluded — no English report: banks publishing sustainability reports only in local languages without an English translation for the relevant year (e.g., RBI 2023 German-only).

Excluded — no standalone ESG/sustainability report: banks producing only annual reports with no dedicated sustainability content meeting a minimum word-count threshold (e.g., ING Group).

Excluded — insufficient coverage: fewer than three usable report-years in the 2020–2024 window (e.g., NJCB: one year only; Nationwide Building Society: only two years of dedicated ESG/sustainability reports).

Excluded — sanctions and data-integrity concerns: VTB Bank, Sberbank (Russia — post-2022 sanctions limit reliable financial data).

The final analytical sample comprises listed commercial banks across five regions (Asia-Pacific, Europe, North America, Latin America, Middle East), yielding the bank-year observations reported in Chapter 5 after merging ESG metrics with financial-market data. A small number of unlisted cooperative and state-owned institutions for which stock-market data is unavailable (BPCE, La Banque Postale, Raiffeisen Switzerland, Crédit Mutuel,

Norinchukin, Rabobank, DZ Bank, LBBW) appear in the ESG-text pipeline but cannot enter the regression panel because the dependent variable (annualised stock return volatility) cannot be computed.

4.1.2 Report Classification

Each bank-year observation is classified into one of three report types:

Type	Code	Description
Standalone ESG/CSR/Sustainability	0	Dedicated non-financial report
Integrated	1	Annual report with embedded sustainability content (IIRC-aligned or CSRD)
Annual report only	3	Excluded from main sample

A **report type dummy** (0 = Standalone, 1 = Integrated) is included as a control variable in all models, as integrated reports may differ systematically in structure and length from standalone ESG reports.

4.1.3 Sample Period

The 2020–2024 window captures: the COVID-19 shock (2020, elevated volatility baseline); the post-pandemic recovery (2021–2022); the CSRD transition period (2023–2024, increasing regulatory pressure on European banks). Year fixed effects absorb these common shocks.

4.2 ESG Report Data Collection

PDF reports were sourced directly from each bank's investor relations or sustainability webpage. Files were standardised using CamelCase folder naming (e.g. JPMChase, BankComms). Text extraction proceeded through a four-stage pipeline:

Standard extraction (PyMuPDF): used for digitally-created PDFs with intact text layers.

Encoding-corrected extraction: applied to PDFs with corrupted font encoding (e.g. BankComms 2020).

Forced OCR (Tesseract 5.0): applied to image-based PDFs (e.g. ChinaEverbrightBank 2020, IndustrialBank 2021).

Language filtering: English-language text identification using word-level frequency matching against English and target-language indicator word lists; non-English text excluded.

Clean word counts range from approximately 4,000 to 661,000 words per report.

4.3 Disclosure Quality Metrics

4.3.1 Fog Index

The Gunning Fog Index measures text readability as a proxy for disclosure complexity:

$$\text{Fog} = 0.4 \times \left(\frac{\text{total words}}{\text{sentences}} + 100 \times \frac{\text{complex words}}{\text{total words}} \right)$$

where complex words are words with three or more syllables (excluding proper nouns, compound words, and common suffixes). Higher Fog scores indicate less readable text. A score of 17 corresponds to college-level reading; most ESG reports in the sample fall in the 18–25 range.

H1 prediction: Higher Fog (less readable) → higher volatility.

4.3.2 L-M Sentiment tone

Using the Loughran-McDonald (2011) financial sentiment dictionary, the net tone of each report is computed as:

$$\text{Tone} = \frac{\text{Positive words} - \text{Negative words}}{\text{Total classified words}}$$

The L-M dictionary is preferred over general-purpose sentiment lexicons (e.g. Harvard General Inquirer) because it was validated specifically on financial texts, avoiding misclassification of financial-domain words (e.g. "liability", "tax" — negative in general but neutral in finance).

H1 prediction: More positive tone (higher Tone) → ambiguous; credibility loss from over-optimism may increase uncertainty. Expected negative coefficient if investors discount overly positive reports.

4.3.3 ESG Vocabulary Density

The proportion of total clean words matching the Baier et al. (2020) ESG Banking Dictionary — a domain-specific lexicon of ESG-related terms for the banking sector:

$$\text{ESG Density} = \frac{\text{ESG dictionary matches}}{\text{total clean words}} \times 100$$

This measures the concentration of sustainability-related content.

4.3.4 SASB Material Topic Coverage

A three-component metric capturing the breadth, balance, and depth of coverage across the six SASB Commercial Banking material topics (Sustainability Accounting Standards Board, 2023; full keyword dictionary in Appendix B): (1) Data Security, (2) Financial Inclusion, (3) ESG Incorporation in Credit Analysis, (4) Financed Emissions, (5) Business Ethics, (6) Systemic Risk Management.

SASB Density (depth):

$$\text{SASB Density} = \frac{\text{weighted SASB keyword hits}}{\text{total clean words}} \times 100$$

SASB HHI Balance (evenness):

$$\text{HHI Balance} = 1 - \sum_{i=1}^6 s_i^2$$

where s_i^2 is topic i 's share of total SASB hits. Higher values indicate more even coverage across all six topics.

SASB Proportional Coverage (breadth):

$$\text{Proportional Coverage} = \frac{\text{topics addressed}}{\text{total topics}} \times 100$$

In the main regressions, SASB Density and HHI Balance enter jointly; Proportional Coverage is used in robustness checks. Due to high collinearity ($r = 0.86$) between SASB Density and ESG Density, they are entered in separate model specifications.

4.4 Financial and Market Data

4.4.1 Dependent Variable: Annualised Stock Return Volatility

Collected via yfinance (Python). For each bank-year, annualised volatility is computed from daily log returns:

$$\sigma = \text{SD}(\ln R_t) \times \sqrt{252}$$

for the calendar year corresponding to the ESG report. Non-trading days are excluded; a minimum of 200 daily observations per year is required. Volatility is winsorised at the 1st and 99th percentiles to reduce outlier influence.

4.4.2 Control Variables

Variable	Source	Notes
Log Market Cap	Third-party financial database; FX-converted to USD	Year-varying year-end USD market cap; year-end FX conversion to USD applied where the bank's listing currency is not the US dollar.
Return on Equity (ROE)	Annual reports (manual collection)	Year-varying for all 80 listed banks across 2020–2024; defined as net income divided by year-end total stockholders' equity.
Beta	Computed (yfinance daily returns vs MSCI ACWI ETF)	Year-varying; OLS slope of bank daily log returns on iShares MSCI ACWI daily log returns, ≥ 100 overlapping trading days per year.
Report Type Dummy	ESG Bank Tracker	0 = Standalone, 1 = Integrated
Log Word Count	Computed from clean text	Controls for report length

Log Total Assets	S&P Global / annual reports	Year-varying USD billions; sourced from each bank's audited primary financial statements at year-end, with year-end FX conversion to USD where source currency is not the US dollar.
Tier 1 ratio	Pillar 3 reports / annual reports	The fully-loaded ratio is used where disclosed, the transitional ratio where the fully-loaded figure is unavailable (held consistent per bank across years), and the standardised-approach ratio for US banks.

The regressions include five bank-year control variables in addition to the fixed effects. Three are year-varying financial controls described in Section 4.4.4: log market capitalisation (bank size), market beta (systematic risk) and return on equity (operating performance). The fourth is the Tier 1 capital ratio, introduced here as a solvency control. The fifth is the report-level integrated-report dummy defined in Section 4.1.2, accompanied by log word count as a report-length control.

The Tier 1 capital ratio is the ratio of a bank's Tier 1 (going-concern) regulatory capital to its risk-weighted assets, and is the supervisory measure of loss-absorbing capital strength under the Basel III framework. It is included because capital adequacy captures a dimension of bank risk that is conceptually distinct from market beta (systematic risk), ROE (profitability) and market capitalisation (size): a better-capitalised bank has a larger buffer against losses and, all else equal, should display lower equity-return volatility. The fully-loaded (Basel III end-state) ratio is used wherever it is disclosed; for the small number of banks that report only the transitional (phased-in) ratio, that measure is used consistently across all five years, and for US banks the standardised-approach ratio is used throughout, because the advanced-approach figure is not reported for every bank-year. Values are hand-collected from primary Pillar 3 reports and audited annual disclosures and entered as decimals; for banks whose fiscal year ends 31 March (the Japanese banks and HDFC Bank) the ratio is taken at the corresponding fiscal year-end, consistent with the dating convention applied to total assets and market capitalisation. Coverage is complete for all 80 panel banks across 2020-2024.

4.4.3 Total Assets (Bank Size)

Total assets are expressed in USD billions throughout the panel and sourced from each bank's audited primary financial statements at year-end, converted to USD at the year-end reference rate from the ECB, HKMA and Bloomberg in the relevant currency (EUR, GBP, CHF, JPY, KRW, CNY, HKD, INR, AED, QAR, AUD, CAD, BRL, DKK, SEK, SGD). For banks whose fiscal year ends 31 March (notably Mizuho Financial Group), the corresponding March year-end balance-sheet figure and 31 March exchange rate are used; for all other banks the figures and rates are taken at 31 December. The S&P Global Top 100 Banks ranking is used solely to define the eligible bank universe for sampling purposes and not as a source of accounting values; this avoids the pro-forma M&A adjustments that S&P applies to certain bank-years and ensures that the size control entering the regression reflects each bank's balance sheet as reported and as priced by the market in the corresponding year. The most consequential single-bank application of this policy is UBS 2022: the S&P 2022 ranking lists UBS at USD 1,679 billion on a pro-forma basis that incorporates the Credit Suisse acquisition (announced March 2023, completed June 2023). The panel uses UBS's reported year-end 2022 total assets of approximately USD 1,104 billion, which is the figure investors observed at the time and which is consistent with the rest of the variable set drawn directly from each bank's own financial statements.

4.4.4 Limitations of Financial Controls

All three primary financial controls — market capitalisation, beta and return on equity — are constructed as fully year-varying variables for the 80 listed banks in the regression sample. Market capitalisation is computed at year-end as the product of year-end shares outstanding and the year-end closing price, converted to USD at the corresponding year-end reference rate where the bank's listing currency is not the US dollar. Beta is computed year-by-year as the slope coefficient from regressing each bank's daily log returns on daily log returns of the iShares MSCI ACWI ETF for the corresponding calendar year, with a minimum of 100 overlapping trading days per year. Return on equity is sourced bank-by-bank from each bank's published audited annual reports across all five years 2020-2024 and is defined throughout as net income divided by year-end total stockholders' equity (the standard textbook ROE), regardless of which equity sub-aggregate any particular bank may emphasise in its own reporting. A fourth control, the Tier 1 capital ratio (Section 4.4.2), is

added in the augmented specification (Model 1b+T1) as a year-varying measure of regulatory capital strength. This control set strengthens the identification of within-bank effects relative to specifications using static snapshots; nevertheless, the short five-year panel still has limited power for bank fixed-effects estimation, a limitation that is examined formally in Section 5.5 through a Mundlak (1978) correlated-random-effects decomposition and a regression-based Hausman test rather than asserted informally.

4.5 Regression Models

4.5.1 Baseline Model (H1)

$$\text{Volatility}_{i,t} = \alpha + \beta_{1\text{Fog}_{i,t}} + \beta_{2\text{Tone}_{i,t}} + \beta_{3\text{SASB_Density}_{i,t}} + \beta_{4\text{HHI}_{i,t}} + \gamma X_{i,t} + \delta_t + \lambda_r + \varepsilon_{i,t}$$

where i indexes banks, t indexes years (2020–2024), the control vector enters with year fixed effects (Asia-Pacific and 2020 omitted) and region fixed effects, and standard errors are clustered at the bank level to account for within-bank serial correlation.

Model specifications:

1. Model 1a: Fog + Tone + ESG Density + SASB HHI (with ROE and Beta controls)
2. Model 1b: Fog + Tone + SASB Density + SASB HHI (with ROE and Beta controls)
3. Model 1b+T1: Model 1b augmented with the Tier 1 capital ratio (the augmented full specification; base for the Section 5.5 / 5.3 extensions)
4. Model 1c: Model 1b without ROE/Beta (the disclosure-quality-only specification, full sample, $N=390$; the preferred cross-sectional test of H1)
5. Model 1d: Individual IV regressions (one metric at a time)
6. Model 1e: SASB Proportional Coverage (alternative SASB metric)

The specifications form a ladder of increasing control richness: Model 1c isolates the cross-sectional disclosure-quality association; Model 1b adds year-varying operating performance (ROE) and systematic risk (beta); and Model 1b+T1 adds the Tier 1 capital ratio. "The full specification" refers to Model 1b+T1 throughout; Model 1b denotes the full-control model without the capital ratio. The purpose of the ladder is to show that the disclosure-quality coefficients are stable as the financial-control set is progressively enriched, rather than to treat any single richer model as superseding the preferred cross-sectional test in Model 1c.

4.5.2 Impression Management Model (H2)

$$\text{Quality Metric}_{i,t} = \alpha + \beta \text{ROE}_{i,t} + \gamma Z_{i,t} + \delta_t + \lambda_r + \varepsilon_{i,t}$$

where the dependent variable is each disclosure quality metric in turn (Fog, Tone, SASB Density, |Tone|), and ROE is the key independent variable. A negative coefficient on ROE in the Fog equation (lower performance → higher Fog) would support H2.

4.5.3 Materiality Premium Model (H3)

Estimated by comparing the R^2 of two nested models:

- Model A: Fog + Tone + ESG Density (generic)
- Model B: Fog + Tone + SASB Density (material)

If H3 holds, Model B should produce systematically higher R^2 and a larger standardised coefficient on SASB Density vs. ESG Density in Model A.

4.5.4 Robustness Checks

Check	Description
R1	Exclude 2020 (COVID shock year)
R2	Flesch-Kincaid Grade instead of Fog Index
R3	Standalone reports only (exclude Integrated)
R4	Asia-Pacific subsample
R5	Add log(Total Assets) as bank size control (N≈390)
RE/FE	Random and entity fixed effects (bank FE, clustered SE)

Robustness is examined along four dimensions. First, conventional sample- and specification-robustness checks address the 2020 COVID-shock baseline (R1), the choice of readability index (R2), the heterogeneity between standalone and integrated reports (R3), regional weighting (R4), the omission of an explicit bank-size control (R5) and the addition of the Tier 1 capital ratio (R6), together with the topic-conditioned reformulation reported in Tables 8 and 9. Second, the robustness of the main results to alternative residual error

structures is examined via Beck-Katz panel-corrected standard errors and feasible GLS estimators (Table 7). Third, the choice between pooled OLS and bank fixed effects is settled formally with a Mundlak correlated-random-effects decomposition and a regression-based Hausman test (Table 10), and the temporal ordering of the disclosure-volatility relationship is examined by re-estimating the model on one-year-lagged disclosure measures (Table 11). Fourth, whether the tone effect is contingent on bank performance is tested with a Tone x ROE interaction (Table 4B), reported in Section 5.3 because it bears directly on the impression-management interpretation.

5. Results

5.1 Descriptive Statistics

The final analytical sample comprises 390 bank-year observations covering 80 listed commercial banks over the period 2020 to 2024. The cross-sectional distribution across regions is: Asia-Pacific 38 banks, Europe 20 banks, North America 17 banks, Latin America 3 banks, Middle East 2 banks. The year distribution is balanced — 76 to 80 banks per year. Standalone ESG, CSR and sustainability reports account for 285 of the 390 report-type observations (73.1%); the remaining 105 are integrated annual reports with embedded sustainability content.

Table 1 reports the summary statistics for the dependent variable, the four disclosure-quality metrics, the financial controls and the report-length control. The dependent variable, annualised stock return volatility, has a mean of 29.5% with a standard deviation of 11.4%, ranging from 12.9% at the first winsorised percentile to 63.9% at the ninety-ninth. This dispersion reflects three sources of cross-sectional and time-series variation: the structural difference between developed-market and emerging-market banking stocks, the elevated baseline of 2020 (the only year in which the post-winsorisation mean exceeds 0.40), and the substantial idiosyncratic variation attributable to bank-specific events such as regulatory penalties, supervisory interventions and the consumer-credit stress at the height of the 2020 COVID shock.

Table 1: Descriptive Statistics

Variable	N	Mean	Std. Dev.	Min	P25	Median	P75	Max
Stock Return Volatility	390	0.2952	0.1136	0.1287	0.2161	0.2710	0.3436	0.6394
Fog Index	390	21.7612	1.9212	17.8576	20.5066	21.4835	22.7357	27.9733
L-M Tone	390	0.3139	0.1988	-0.1286	0.1902	0.3455	0.4612	0.6946
ESG Vocabulary Density (%)	390	7.0563	1.9494	2.3421	5.5705	7.1848	8.5527	11.1020
SASB Density (%)	390	2.4959	0.8958	0.9614	1.7816	2.4498	3.1426	4.9604
SASB HHI Balance	390	0.8860	0.0927	0.5335	0.8600	0.9142	0.9482	0.9801
SASB Proportional Coverage (%)	390	86.20	18.07	16.67	83.33	100.00	100.00	100.00
Log Market Cap	390	24.5011	0.8730	22.7730	23.8977	24.4951	24.9705	26.8737
Return on Equity	390	0.1038	0.0438	-0.0262	0.0803	0.1011	0.1300	0.2074
Beta	390	0.5386	0.4065	-0.0976	0.2017	0.4638	0.8119	1.5262
Tier 1 Capital Ratio	390	0.1479	0.0262	0.0978	0.1320	0.1480	0.1640	0.2199
Integrated Report (dummy)	390	0.2692	0.4441	0.0000	0.0000	0.0000	1.0000	1.0000
Log Word Count	390	10.9615	0.8014	9.2431	10.4168	10.8316	11.4141	12.9682
Word Count	390	84,099	92,875	4,035	33,417	50,597	90,599	660,979

Notes: Sample consists of 80 listed commercial banks, 2020–2024 (N = 390 bank-year observations). All continuous variables winsorised at the 1st and 99th percentiles. ROE for AgriCultural Bank of China, JPMorgan Chase, Goldman Sachs and Bank of New York Mellon is recalculated as net income divided by year-end total stockholders' equity rather than the common or tangible equity figures reported by third-party data providers. Total assets are sourced directly from primary annual reports with year-end currency conversion to USD, replacing pro-forma S&P-ranking values for eight banks. The Capital One 2020 ESG report is dropped from the regression panel as a combined 2019–2020 disclosure straddling the sampling window. Tier 1 Capital Ratio is the bank's fully-loaded Tier 1 capital ratio, sourced from primary Pillar 3 and annual-report disclosures and expressed as a decimal.

The mean Fog Index is 21.76 with a standard deviation of 1.92, placing the average ESG report at roughly early-college reading difficulty. The interquartile range of 20.51 to 22.74 indicates that ESG-report readability is relatively narrowly distributed within the sample, while the tails extend from clearly accessible reports (Fog < 18) to text approaching graduate-level complexity (Fog > 27). Mean L-M Tone is 0.314, indicating that across the sample positive financial-sentiment words outnumber negative ones by approximately two to one — consistent with the documented tendency of ESG reports to adopt aspirational rather than risk-disclosure-oriented framing. Tone exhibits considerable cross-sectional variation (SD = 0.199; range –0.129 to 0.695 after winsorisation), with the lower tail occupied by reports that engage more candidly with adverse outcomes — regulatory enforcement actions, climate-related credit losses, litigation exposures and, in the case of Société Générale's 2024 Universal Registration Document, the absorption of an annual-report editorial register that produces a markedly negative net positive sentiment.

ESG vocabulary density (mean 7.06%, SD 1.95%) is substantially higher than SASB material-topic density (mean 2.50%, SD 0.90%), confirming that even in heavily sustainability-themed reports the proportion of text devoted specifically to the six SASB Commercial Banking material topics is comparatively modest. SASB HHI Balance — a measure of how evenly coverage is distributed across the six topics — has a mean of 0.886 (maximum 0.980), implying that the typical bank covers all six topics with substantial residual variation across institutions. SASB Proportional Coverage averages 86.2%, meaning the typical bank addresses approximately five of the six material topics, with a median of 100% (full coverage). These three SASB components capture distinct dimensions of disclosure: density measures depth, HHI Balance measures evenness, and Proportional Coverage measures breadth.

The financial controls span a wide range. Log market capitalisation, constructed from year-end USD market caps collected from primary annual reports and third-party financial databases, has a mean of 24.50 with a standard deviation of 0.87, corresponding to a USD

market cap range of approximately USD 8 billion to USD 480 billion. Return on equity, constructed as a fully year-varying variable for all 80 banks in the sample and corrected on a bank-by-bank basis where source data reported return on common or tangible equity rather than total-equity ROE, has a mean of 10.38% with a standard deviation of 4.38%, ranging from -2.6% at the first winsorised percentile to 20.7% at the ninety-ninth. The dispersion captures both genuine cross-bank heterogeneity (state-supported Chinese banks at the lower end of profitability, Latin American and Italian banks at the upper end) and time-series fluctuation associated with the 2020 COVID shock and the subsequent recovery. Beta is also constructed as a fully year-varying variable, computed as the slope coefficient from regressions of each bank's daily log returns on daily log returns of the iShares MSCI ACWI ETF for the corresponding calendar year. The sample mean is 0.54 with a standard deviation of 0.41, range -0.10 to 1.53 (winsorised); beta values are highest in 2020 reflecting elevated systematic risk during the COVID shock.

The Tier 1 capital ratio, included as a solvency control in the augmented specification, has a sample mean of 14.8% with a standard deviation of 2.6 percentage points and ranges from approximately 9.8% to 22.0% after winsorisation. The modest cross-sectional dispersion reflects the composition of the sample — large, well-capitalised institutions operating under broadly comparable Basel III minimum requirements — and foreshadows the limited marginal explanatory power of the capital ratio once market risk and profitability are controlled for (Section 5.2).

Table 2 reports the bivariate correlation matrix. All four disclosure-quality metrics exhibit weak negative pairwise correlations with annualised volatility: Fog ($r = -0.13$), L-M Tone ($r = -0.14$), ESG Density ($r = -0.12$) and SASB Density ($r = -0.14$). The Tone correlation is meaningfully stronger in the final sample than in earlier iterations of this study, an early indication of the strengthening of the H1b finding documented in Section 5.2. Second, ESG Density and SASB Density are very strongly correlated ($r = 0.854$), confirming the collinearity concern flagged in the methodology; this correlation underpins the decision to enter these two variables in separate model specifications and is central to the failure of the materiality premium hypothesis examined in Section 5.4. Third, the integrated-report dummy exhibits substantial negative correlations with both ESG Density ($r = -0.69$) and SASB Density ($r = -0.58$) and with L-M Tone ($r = -0.49$), reflecting the structural difference between dedicated sustainability reports and annual reports with embedded sustainability

content. Fourth, ROE — fully year-varying and now calculated on the standard total-equity basis for all 80 banks — is strongly negatively correlated with volatility ($r = -0.39$). Log market capitalisation, denominated in USD, is negatively correlated with volatility ($r = -0.18$), aligning with the standard size–risk relationship in banking. Beta, computed against the MSCI ACWI benchmark, is strongly positively correlated with volatility ($r = 0.60$), reflecting the mechanical relationship between beta and volatility.

Table 2: Pairwise Correlation Matrix

	Vol	Fog	Tone	ESG%	SASB %	HHI	lnMC	ROE	Beta	T1	IntRpt
Vol	1.000										
Fog	-0.134* **	1.000									
Tone	-0.136* **	0.153** *	1.000								
ESG %	-0.124* *	0.286** *	0.318** *	1.000							
SASB %	-0.141* **	0.291** *	0.110** *	0.854** *	1.000						
HHI	-0.022	0.061	0.234** *	0.164** *	0.034	1.000					
lnMC	-0.179* **	-0.100* *	0.166** *	0.068	0.002	-0.046	1.000				
ROE	-0.392* **	-0.098* *	0.096* *	0.169** *	0.084* *	0.182* **	0.310** *	1.000			
Beta	0.595** *	-0.283* **	-0.031	0.032	-0.025	-0.053	0.181** *	-0.128* **	1.000		
T1	0.109**	-0.246* **	-0.217* **	-0.043	0.055	-0.088 *	0.069	0.038	0.228* **	1.000	
IntRpt	0.136** *	-0.351* **	-0.487* **	-0.687* **	-0.580* **	-0.121 **	-0.182* **	-0.069	0.039	0.236* **	1.000

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Pairwise complete observations ($N = 390$). Vol = Annualised volatility; Fog = Fog Index; Tone = L-M Tone; ESG% = ESG Vocabulary Density; SASB% = SASB Density; HHI = SASB HHI Balance; lnMC = Log Market Cap; T1 = Tier 1 Capital Ratio; IntRpt = Integrated Report dummy.

A topic-conditioned analysis of the textual material confirms a striking descriptive pattern that motivates several of the robustness checks reported in Section 5.5. In each of the 390 bank-year reports, sentences containing at least one SASB primary-keyword match are flagged as "material" and the remainder as "non-material". Across the sample, material content accounts for only 12% of sentences and 19% of words. Within each report, the Fog Index of material sentences exceeds the Fog Index of non-material sentences by an average of 9.54 points (mean Fog of material content = 30.1; mean Fog of non-material content = 20.6); and the Loughran-McDonald tone of material content is less positive than the tone of non-material content by an average of 0.27 (mean Tone of material content = +0.10; mean Tone of non-material content = +0.37). Both gaps are extremely well-determined statistically and consistent across the sample: the Fog gap is positive in 388 of 390 reports (99.5%) and the Tone gap is negative in 356 of 390 reports (91.3%); the corresponding

Cohen's d effect sizes are +1.76 and -1.43 , both in the very-large category. Table 8 reports the formal tests, including per-year and per-report-type breakdowns confirming that the pattern is structural rather than driven by any particular year or document format. The substantive interpretation of this descriptive finding, and the regression evidence that it operates as an industry-wide level effect rather than as a discriminator across banks, are taken up in Section 5.5 and Section 6.1.

Variance inflation factor diagnostics, computed on the five core independent variables, indicate that no individual coefficient is materially destabilised by collinearity when ESG and SASB densities are not jointly included. With all five disclosure metrics in the model, ESG Density and SASB Density yield VIFs of 4.7 and 4.3 respectively — high but below conventional thresholds — and dropping either variable returns all VIFs to comfortably acceptable levels (≤ 1.30), reinforcing the case for the model-pair approach adopted in this thesis.

The Tier 1 capital ratio is added to the correlation matrix in Table 2. Its pairwise correlation with volatility is positive but weak ($r = +0.11$), the opposite sign to the partial effect it takes once systematic risk is controlled for; this reversal arises because higher-capital banks in the sample also tend to carry higher beta ($r = +0.23$) and to publish integrated rather than standalone reports ($r = +0.24$). The capital ratio is essentially uncorrelated with ROE ($r = +0.04$) and bank size ($r = +0.07$), confirming that it captures a distinct dimension of bank risk; its variance inflation factor in the augmented specification is 1.11.

5.2 H1 – Disclosure Quality and Stock Return Volatility

Table 3: OLS Regression Results — Disclosure Quality and Stock Return Volatility (H1)

Dependent Variable: Annualised Stock Return Volatility

Variable	(1)	(2)	(3)	(4)	(5)	(6)
Fog Index	0.0067** (0.0034)		0.0072** (0.0032)	0.0072** (0.0031)	0.0034* (0.0019)	0.0033* (0.0019)
L-M Tone		-0.1198^{***} (0.0364)	-0.1199^{***} (0.0357)	-0.1223^{***} (0.0374)	-0.0550^{**} (0.0254)	-0.0566^{**} (0.0250)
ESG Vocabulary Density			0.0024 (0.0045)			
SASB Density				0.0034 (0.0089)	0.0029 (0.0063)	0.0046 (0.0062)

SASB HHI Balance				0.0049 (0.0465)	0.0512* (0.0310)	0.0482 (0.0308)
Log Market Cap	−0.0114 (0.0089)	−0.0125 (0.0086)	−0.0070 (0.0087)	−0.0068 (0.0088)	−0.0148*** (0.0055)	−0.0136** (0.0054)
ROE					−0.2058*** (0.0785)	−0.1948** (0.0765)
Beta					0.2300*** (0.0153)	0.2306*** (0.0154)
Tier 1 Capital Ratio						−0.1943 (0.1675)
Integrated Report	0.0427*** (0.0143)	0.0190 (0.0146)	0.0339** (0.0170)	0.0330** (0.0164)	0.0229** (0.0100)	0.0260*** (0.0099)
Log Word Count	−0.0114 (0.0079)	−0.0218** (0.0086)	−0.0253*** (0.0085)	−0.0253*** (0.0084)	−0.0093* (0.0050)	−0.0099** (0.0049)
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Region Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	390	390	390	390	390	390
R-squared	0.5058	0.5197	0.5305	0.5305	0.7894	0.7899
Adjusted R-squared	0.4905	0.5050	0.5117	0.5117	0.7793	0.7797
F-statistic	50.20	44.62	40.31	36.58	151.78	163.67

Notes: Standard errors clustered at the bank level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Year fixed effects (2021–2024 dummies, base = 2020) and region fixed effects (Europe, Latin America, Middle East, North America dummies, base = Asia-Pacific) are included in all specifications but the coefficients are suppressed for brevity. Sample: 80 listed commercial banks, 2020–2024 ($N = 390$). Specifications (1) and (2) are the univariate Model 1d regressions of volatility on Fog Index and L-M Tone respectively, each with base controls. Specification (3) is the Model 1c specification using ESG Vocabulary Density. Specification (4) is the preferred Model 1c specification using SASB Density together with SASB HHI Balance. Specification (5) is Model 1b, which adds year-varying ROE and beta. Specification (6) augments Model 1b with the bank's fully-loaded Tier 1 capital ratio — the thesis's final control specification — under which the Fog Index and L-M Tone coefficients are essentially unchanged.

The main regression results are reported in Table 3. Specifications (1) and (2) are the univariate Model 1d regressions of volatility on the Fog Index and on L-M Tone respectively, each with base controls. Specifications (3) and (4) are the Model 1c specifications with all three stylistic dimensions jointly entered — (3) using ESG vocabulary density and (4) using SASB Density together with SASB HHI Balance. Specification (5) is Model 1b, which augments (4) with year-varying ROE and year-varying beta. Specification (4) — the preferred cross-sectional test of H1 — corresponds to the disclosure-quality-only specification that does not condition on the systematic-risk and operating-performance channels. Specification (6) augments Model 1b with the bank's fully-loaded Tier 1 capital ratio.

The headline findings are that L-M Tone is significantly negatively associated with annualised stock return volatility, that the Fog Index displays a positive association consistent with the obfuscation channel, and that bank size (denominated in USD) enters with the theoretically expected negative sign when the full set of financial controls is included. In Model 1c (column 4, $N = 390$), the tone coefficient is -0.122 (clustered standard error 0.037 , $p = 0.001$): a one-standard-deviation increase in L-M Tone (approximately 0.20 units) is associated with a 2.4 percentage point reduction in annualised volatility, equivalent to roughly 21% of one cross-sectional standard deviation in the dependent variable. The economic magnitude is substantial when set against the average volatility level of 29.5% . In the univariate regression (Model 1d, column 2), the tone effect is essentially identical ($\beta = -0.120$, $p = 0.001$), confirming that the result is not driven by interaction with the other disclosure metrics. The standalone-reports subsample (R3, Section 5.5) yields a substantially stronger effect ($\beta = -0.203$, $p < 0.001$), and the bank-size-controlled robustness check (R5) yields $\beta = -0.130$ at $p = 0.001$.

The Fog Index coefficient in Model 1c is positive and significant at the 5% level: $\beta = +0.0072$ (clustered standard error 0.0031 , $p = 0.022$). A one-unit increase in Fog Index (approximately one half of a standard deviation) is associated with a 0.7 percentage point increase in annualised volatility. The point estimate is smaller than the tone effect but theoretically important: its sign matches the obfuscation-channel prediction of H1a, and the statistical significance of the Fog effect strengthens noticeably relative to earlier iterations of this analysis following the pre-submission corrections to ROE, the replacement of pro-forma S&P-sourced total-asset values with report-sourced year-end values, and the exclusion of non-universe banks. The Fog effect remains significant at the 5% level in the univariate Model 1d ($\beta = +0.0067$, $p = 0.048$) and in the standalone-reports subsample (R3, $\beta = +0.0068$, $p = 0.029$), attenuates to marginal- 10% significance in the full-control Models 1a/1b ($\beta = +0.0034$, $p = 0.085$ in Model 1a; $\beta = +0.0034$, $p = 0.078$ in Model 1b), and is significant under all four alternative error-structure estimators applied to Model 1c — at the 1% level under Beck-Katz PCSE and FGLS with panel heteroskedasticity, the 5% level under OLS with clustered standard errors, and the 10% threshold under FGLS with AR(1) errors ($p = 0.088$). The persistence of the Fog point estimate across specifications, together with the strongly significant impression-management result reported in Section 5.3 (ROE $\rightarrow$ Fog at $p = 0.028$), provides convergent evidence for the readability channel.

Log market capitalisation enters Model 1c with the theoretically expected negative sign and is insignificant in the disclosure-quality-only specification ($\beta = -0.007$, $p = 0.437$), and strengthens to $\beta = -0.015$ ($p \approx 0.007$) in Models 1a and 1b once the full set of year-varying financial controls is included. A one-standard-deviation increase in log market cap is associated with a 1.3 percentage point reduction in annualised volatility in the full-control specification.

In sharp contrast, neither SASB Density ($\beta = +0.003$, $p = 0.703$ in Model 1c) nor SASB HHI Balance ($\beta = +0.005$, $p = 0.917$) is significantly associated with stock return volatility. The univariate regressions reported in Model 1d reinforce these conclusions: ESG Density ($\beta = +0.005$, $p = 0.365$), SASB Density ($\beta = +0.012$, $p = 0.236$) and SASB HHI Balance ($\beta = -0.047$, $p = 0.343$), entered individually, are statistically indistinguishable from zero. The materiality dimension thus does not contribute meaningful explanatory power once the linguistic credibility and accessibility dimensions are accounted for. The implications of this null result for the materiality premium hypothesis (H3) are developed in Section 5.4.

Models 1a and 1b — which include the full set of controls together with year-varying ROE and year-varying beta — paint a complementary picture. The model fit jumps substantially relative to Model 1c ($R^2 = 0.79$ in both Models 1a and 1b, compared with $R^2 = 0.53$ in Model 1c). This improvement is driven almost entirely by the beta coefficient: $\beta = +0.230$ in both Model 1a and Model 1b (clustered standard error 0.015, $z \approx 15$, $p < 0.001$). The economic interpretation is straightforward — beta and volatility are mechanically related, both being functions of the same daily-return time series, and a year-matched beta estimate captures the systematic-risk component of volatility directly. Log market capitalisation enters with the expected negative sign and is highly significant ($\beta = -0.015$, $p = 0.006$). ROE, consistently calculated as net income over end-of-period total stockholders' equity, retains a strongly negative and significant coefficient ($\beta = -0.209$, $p = 0.008$ in Model 1a; $\beta = -0.206$, $p = 0.009$ in Model 1b). A one-standard-deviation increase in ROE (approximately 4.4 percentage points) is associated with a 0.9 percentage point reduction in annualised volatility, holding constant the systematic-risk channel captured by beta.

In Models 1a and 1b, the L-M Tone coefficient remains significant at the 5% level ($\beta = -0.057$, $p = 0.027$ in Model 1a; $\beta = -0.055$, $p = 0.030$ in Model 1b). This is a clear strengthening relative to earlier iterations of this analysis with fewer data points, in which the tone coefficient attenuated to marginal significance once beta was included; the revised

analysis indicates that the cross-sectional tone effect survives the inclusion of year-matched systematic risk and operating performance. The Fog Index coefficient retains marginal-10% significance: $\beta = +0.0034$ ($p = 0.085$) in Model 1a and $\beta = +0.0034$ ($p = 0.078$) in Model 1b. The retention of both stylistic effects at conventional or near-conventional significance levels, even after partialling out year-matched systematic risk through beta and operating performance through ROE, sharpens the H1 interpretation: tone and readability are not merely correlates of financial fundamentals but operate as independent disclosure-quality channels.

Column (6) of Table 3 reports the augmented full specification, Model 1b+T1, which adds the Tier 1 capital ratio to Model 1b. The capital ratio enters with the theoretically expected negative sign — better-capitalised banks display lower equity-return volatility — but the coefficient is not statistically significant (beta = -0.194, clustered standard error 0.168, $p = 0.246$). The weak marginal effect is unsurprising: capital ratios are tightly clustered across these large banks (Section 5.1), and the systematic-risk and operating-performance channels captured by beta and ROE already absorb most of the cross-sectional variation in volatility. The result of substantive interest is that the two disclosure-quality coefficients are essentially unchanged by the inclusion of the solvency control — the Fog Index coefficient moves from +0.0034 to +0.0033 ($p = 0.085$) and the L-M Tone coefficient from -0.0550 to -0.0566 ($p = 0.024$), each retaining its sign, magnitude and significance. The headline H1 findings are therefore robust to controlling for regulatory capital, and Model 1b+T1 serves as the base specification for the estimator-choice, temporal-ordering and performance-moderation analyses reported in Sections 5.5 and 5.3.

Taken together, the Model 1 results yield three conclusions for H1. First, H1b (the tone channel) is supported in the cross-sectional specification (Model 1c, $\beta = -0.122$, $p = 0.001$) and across every robustness check that retains adequate cross-sectional identifying variation. Second, H1a (the readability channel) is supported at the 5% level in Model 1c, in the univariate Model 1d and in the standalone-reports subsample (R3), and at the 10% level in the full-control Models 1a/1b and under each of the four alternative error-structure estimators. Third, H1c (the ESG density and SASB coverage channels) is not supported in any cross-sectional specification examined. Hypothesis H1 is therefore strongly supported in its tone and readability components and not supported in its substantive-content components.

5.3 H2 – Impression Management and Financial Performance

The impression management hypothesis (H2) predicts that financially weaker banks strategically alter the linguistic style of their ESG disclosures — either by increasing complexity (obfuscation) or by inflating positive tone — to manage external perceptions of performance. Each of the disclosure-quality metrics is regressed in turn on ROE, with the full set of financial controls (excluding the dependent-variable metric itself), report-length, and year and region fixed effects. The estimates are reported in Table 4.

Table 4: OLS Regression Results — Impression Management (H2)

Variable	H2a: Fog Index	H2b: L-M Tone	H2c: SASB Density	H2 additional: Tone
ROE	−5.5464** (2.5063)	0.3660 (0.2691)	−0.9557 (1.0739)	0.1786 (0.2438)
Log Market Cap	−0.2588 (0.1903)	0.0246 (0.0179)	−0.1611** (0.0644)	0.0050 (0.0184)
Integrated Report	−1.7244*** (0.4126)	−0.0571 (0.0510)	−1.1973*** (0.1606)	−0.0388 (0.0512)
Log Word Count	0.1888 (0.2742)	−0.1554*** (0.0285)	0.0379 (0.0789)	−0.0316 (0.0298)
Year Fixed Effects	Yes	Yes	Yes	Yes
Region Fixed Effects	Yes	Yes	Yes	Yes
Observations	390	390	390	390
R-squared	0.3408	0.4950	0.5163	0.4467
Adjusted R-squared	0.3195	0.4787	0.5005	0.4287

*Notes: Standard errors clustered at the bank level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Sample: 80 listed commercial banks, 2020–2024 ($N = 390$). Each column is a separate regression with the indicated disclosure-quality metric as the dependent variable and ROE as the key independent variable. H2 predicts: weaker performance (lower ROE) → less readable reports (higher Fog) and more optimistic tone (higher Tone). Year and region fixed effects included in all columns.*

The strongest result emerges in the Fog Index equation. The coefficient on ROE is −5.55 (clustered standard error 2.51, $p = 0.028$), indicating that a one-standard-deviation decrease in ROE (4.4 percentage points) is associated with an increase in the Fog Index of approximately 0.24 units. The economic magnitude is meaningful — approximately one eighth of one cross-sectional standard deviation in readability — and the statistical significance is at the conventional 5% threshold. The direction of the effect aligns squarely with the classical obfuscation literature (Courtis, 1998; Lo, Ramos & Rogo, 2017) extended to the ESG-disclosure setting: when financial performance is weak, managers produce systematically less readable sustainability reports.

The tone equation does not support the impression-management prediction. The coefficient on ROE is positive but not significantly different from zero ($\beta = +0.366$, $p = 0.175$). The data therefore reject the tone-inflation channel of impression management: weaker-performing banks do not produce more positively toned reports than stronger-performing banks. The absolute-tone regression confirms that the extremity of tone is also not significantly related to performance ($\beta = +0.18$, $p \approx 0.45$), ruling out a non-monotonic relationship in which the most extreme tones systematically occur in either high- or low-performing banks.

The SASB density equation shows no significant relationship between financial performance and disclosure content. The coefficient on ROE is -0.98 ($p = 0.372$), well outside conventional significance bounds. In the absence of significance for the tone-inflation and SASB-density channels, the substantive content of ESG disclosure appears to respond less to short-term financial performance than the stylistic dimensions of disclosure — a pattern consistent with disclosure frameworks (SASB topic coverage, governance disclosures) being relatively sticky from year to year, while stylistic choices (readability, tone) are made more flexibly in light of contemporaneous performance.

Several control variables in the H2 regressions are noteworthy. The report-type dummy enters strongly and negatively in the Fog equation ($\beta = -1.74$, $p < 0.001$): integrated reports are systematically less complex than standalone ESG reports, attributable to the broader audience and editorial discipline of integrated annual reports. In the tone equation, log word count is strongly negative ($\beta = -0.156$, $p < 0.001$), confirming that longer reports — which tend to engage more substantively with adverse risks — exhibit lower net positive sentiment.

Taken together, the H2 results provide direct support for the classical obfuscation hypothesis (lower performance $\rightarrow$ less readable reports, $p = 0.028$) and do not support the tone-inflation hypothesis (no significant relationship between performance and tone). This asymmetry is itself an important finding. The two stylistic channels of impression management operate independently: managers respond to weak performance by hiding bad news in complex prose (obfuscation), but the data offer no evidence that they then spin it with optimistic language. The textual style of ESG reports thus carries two distinct signals for investors — readability tracks contemporaneous financial performance, while tone does not — and the literature should treat the two as separate strategic dimensions rather than twin

manifestations of a single impression-management impulse. H2 is therefore partially supported: obfuscation operates as predicted, tone inflation does not.

Performance-contingency of the tone effect. The direct test of tone inflation (H2b) finds no average relationship between performance and tone: weak performers do not, on average, adopt more positive language. A complementary and distinct question is whether the market's response to tone depends on performance — that is, whether the tone-volatility relationship documented for H1b is itself moderated by ROE. This is tested by augmenting the full specification (Model 1b+T1) with an interaction between L-M Tone and ROE, mean-centring both variables so that each main effect is evaluated at the sample mean of the other. The estimates are reported in Table 4B. The interaction is positive and significant ($\beta = 0.717$, $p = 0.039$): the volatility-reducing association of positive tone weakens as bank performance improves. The marginal-effect panel of Table 4B makes the pattern concrete. For weak performers — one standard deviation of ROE below the mean — a one-unit increase in tone is associated with a 0.086 reduction in annualised volatility ($p = 0.007$); at mean ROE the effect is -0.055 ($p = 0.025$); and for strong performers — one standard deviation above the mean — the effect falls to a statistically insignificant -0.023 ($p = 0.353$). In other words, report tone moves volatility only where impression-management incentives are strongest; for well-performing banks, whose fundamentals are largely self-evident, tone carries no incremental association with risk.

Table 4B: Tone $\times$ Performance Interaction — Impression-Management Moderation

Dependent Variable: Annualised Stock Return Volatility

	Volatility (year t)
Model coefficients (Tone and ROE mean-centred)	
L-M Tone (at mean ROE)	-0.0546^{**} (0.0244)
ROE (at mean Tone)	-0.1732^{**} (0.0691)
Tone $\times$ ROE	0.7172^{**} (0.3483)
Marginal effect of L-M Tone on volatility	
at low ROE (mean -1 SD)	-0.0860^{***} (0.0321)
at mean ROE	-0.0546^{**} (0.0244)
at high ROE (mean $+1$ SD)	-0.0232 (0.0250)
Observations	390
R-squared	0.7924

*Notes: Standard errors clustered at the bank level, in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. The specification is Model 1b augmented with the Tier 1 capital ratio and a Tone $\times$ ROE interaction; L-M Tone and ROE are mean-centred so that each main effect is evaluated at the sample mean of the other. The positive and significant interaction (0.7172) implies that the volatility-reducing association of positive report tone weakens as bank performance improves. The lower panel reports the marginal effect of tone on volatility at low, mean and high ROE, where low and high denote one standard deviation below and above the mean (ROE standard deviation = 0.044): tone is strongly and significantly associated with lower volatility for weak performers (-0.0860 at low ROE) but is economically small and statistically insignificant for strong performers (-0.0232 at high ROE). The tone–volatility relationship is thus concentrated among the banks with the*

greatest incentive to manage impressions, consistent with the impression-management interpretation of H2. Year and region fixed effects and the full control set (log market capitalisation, beta, Tier 1 capital ratio, integrated-report dummy and log word count) are included. Sample: 80 listed commercial banks, 2020–2024 (N = 390).

This performance-contingency does not contradict the H2b null. Weak performers do not systematically inflate their tone; but where a weak performer does adopt positive tone, that tone is associated with materially lower volatility, whereas for a strong performer the same stylistic choice is uninformative. The natural reading is that tone carries the greatest uncertainty-reducing value precisely where investor uncertainty about the bank is greatest. Whether this reflects markets being genuinely reassured by informative tone or instead being placated by strategic language cannot be resolved with these data; the robust empirical regularity is that the tone-volatility link is conditional on performance, which strengthens the case for treating tone as a performance-contingent disclosure signal rather than a uniform one. The interpretive implications are developed in Section 6.3.

5.4 H3 – The Materiality Premium

Hypothesis H3 predicts that disclosure of SASB-defined material topics should explain a greater share of cross-sectional variation in stock return volatility than generic ESG vocabulary, reflecting investors' preferential attention to material content. The hypothesis is tested by comparing two nested specifications reported in Table 5: Model A, which includes Fog, Tone and ESG Vocabulary Density along with controls and fixed effects; and Model B, which is identical except that SASB Density replaces ESG Density.

The two models yield virtually identical fit. Model A produces $R^2 = 0.5305$ (adjusted $R^2 = 0.5117$); Model B produces $R^2 = 0.5305$ (adjusted $R^2 = 0.5117$). The Akaike and Bayesian Information Criteria are similarly close: Model A $AIC \approx -854$, Model B $AIC \approx -854$. The standardised coefficients on the two density measures — designed to enable direct comparison — are both insignificantly different from zero: ESG Density $\beta^* = +0.040$ ($p = 0.592$), SASB Density $\beta^* = +0.027$ ($p = 0.704$). A specification estimating both metrics jointly (Model C) yields an R^2 gain over either single-metric specification of less than 0.0001, with neither density coefficient significant.

The materiality premium hypothesis is therefore not supported. SASB material topic coverage does not explain more variance in volatility than generic ESG vocabulary, and the

standardised coefficient on SASB Density is in fact marginally smaller than that on ESG Density.

Table 5: Materiality Premium Comparison (H3)

Comparing model fit and standardised coefficients of generic ESG density vs SASB material-topic density

Panel A: Model fit

Specification	R ²	Adj. R ²	N	AIC	BIC
Model A: Fog + Tone + ESG Density	0.5305	0.5117	390	-854.0	-794.5
Model B: Fog + Tone + SASB Density	0.5305	0.5117	390	-854.0	-794.5
Model C: Fog + Tone + ESG + SASB (both)	0.5305	0.5104	390	-852.0	-788.5

Panel B: Standardised coefficients (β^*) on density measures

Specification	Coefficient (β)	Std. β^*	p-value
Model A — ESG Vocabulary Density	+0.0024	+0.0404	0.592
Model B — SASB Material Density	+0.0034	+0.0267	0.704

Notes: All models include Fog Index, L-M Tone, log market capitalisation, the integrated-report dummy, log word count, year fixed effects and region fixed effects. Standard errors clustered at the bank level. Sample: 80 listed commercial banks, 2020–2024 (N = 390). H3 predicts that the SASB-density specification (Model B) achieves a higher R² and a larger standardised coefficient than the ESG-density specification (Model A). The evidence is null: $\Delta R^2 < 0.0001$, $\Delta AIC < 0.1$, and the standardised coefficient on SASB density (0.027) is marginally smaller than that on ESG density (0.040), with neither significant. The very high pairwise correlation between ESG and SASB densities ($r = 0.854$; see Table 2) limits separate identification.

Two interpretations of this null result are admissible. The first is methodological: SASB Density and ESG Density are highly collinear ($r = 0.854$). In practice, banks that cover ESG topics extensively also tend to cover SASB material topics extensively. This collinearity reflects the empirical reality that the SASB Commercial Banking material topics are a structured subset of the broader ESG vocabulary, and the two are not separately identifiable in the cross-section without a more refined measurement strategy that isolates material-only content. The robustness analysis reported in Section 5.5 — which constructs Fog and L-M Tone measures from the material-text subset of each report — confirms that the cross-sectional null on the materiality channel reflects a deeper structural pattern in the data: while the within-document level differences in disclosure quality between material and non-material content are very large and highly significant, the cross-bank variation in those gaps does not predict stock-return volatility once the whole-document measures are included.

The second interpretation is substantive: investors may not, in fact, attend differently to material versus generic ESG disclosure in the banking sector at the level of textual coverage. While Khan, Serafeim & Yoon (2016) documented a materiality premium for SASB-defined performance measures, the present results suggest that the equivalent premium for

SASB-defined disclosure is not detectable in this sample. This may reflect the well-documented difficulty investors face in extracting actionable information from qualitative sustainability narratives; the meaningful differentiation may occur at the level of disclosed quantitative data (greenhouse gas footprints, financed emissions, scenario-analysis results) rather than at the level of textual coverage. The robustness check using SASB Proportional Coverage (Model 1e) produces essentially the same null result, reinforcing the conclusion that the materiality dimension as currently measured does not independently shape investor uncertainty.

5.5 Robustness

Table 6: Robustness Checks

Dependent Variable: Annualised Stock Return Volatility

Specification	Fog Index	L-M Tone	SASB Density	SASB HHI	N	R ²
R1: Exclude 2020 (COVID)	+0.0077**	−0.0984**	−0.0043	−0.0158	314	0.209
R2: Flesch-Kincaid (vs Fog)	+0.0049*	−0.1193***	+0.0046	+0.0082	390	0.527
R3: Standalone reports only	+0.0068**	−0.2030***	−0.0002	+0.0699	285	0.500
R4: Asia-Pacific only	+0.0051	−0.1204**	+0.0019	+0.0028	190	0.310
R5: + Log Total Assets	+0.0064*	−0.1296***	+0.0042	+0.0109	390	0.533
Bank Fixed Effects	+0.0055	−0.0121	+0.0288**	+0.1037	390	0.603
Random Effects	+0.0004	−0.0723*	+0.0181**	+0.0295	390	0.557

*Notes: Standard errors clustered at the bank level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. All specifications include log market capitalisation, the integrated-report dummy, log word count, year fixed effects and region fixed effects; coefficients suppressed for brevity. R1 excludes 2020 observations to address the COVID-shock concern. R2 substitutes the Flesch-Kincaid Grade level for the Fog Index. R3 restricts the sample to standalone ESG/sustainability/CSR reports. R4 restricts the sample to Asia-Pacific banks. R5 adds the log of total assets as an explicit bank-size control. Bank Fixed Effects uses a within-bank estimator (region dummies absorbed); the F-test for poolability rejects pooling at the 1% level ($F = 4.61$, $p < 0.001$). Random Effects serves as a Hausman comparison.*

The robustness of the main results is examined along four dimensions. First, conventional sample- and specification-robustness checks are reported in Table 6 — excluding the 2020 COVID year (R1), substituting the Flesch-Kincaid grade for the Fog Index (R2), restricting to standalone reports (R3) and to Asia-Pacific banks (R4), adding an explicit bank-size control (R5) and adding the Tier 1 capital ratio (R6) — together with the topic-conditioned reformulation of the disclosure-quality metrics reported in Tables 8 and 9. Second, the sensitivity of the headline findings to alternative residual error structures is examined via Beck-Katz panel-corrected standard errors and feasible GLS estimators (Table 7). Third, the

choice between pooled OLS and bank fixed effects is settled formally with a Mundlak (1978) correlated-random-effects decomposition and a regression-based Hausman test (Table 10), and the temporal ordering of the disclosure–volatility relationship is examined by re-estimating the model on one-year-lagged disclosure measures (Table 11). The fourth dimension — whether the tone effect is contingent on bank performance (a Tone $\times$ ROE interaction) — is reported in Section 5.3, where it bears directly on the impression-management interpretation.

Table 7: Alternative Error-Structure Estimators on Model 1c

Dependent Variable: Annualised Stock Return Volatility

Variable	OLS + Cluster SE	Beck-Katz PCSE	FGLS AR(1)	FGLS Het.
Fog Index	0.0072** (0.0031)	0.0072*** (0.0018)	0.0050* (0.0030)	0.0079*** (0.0020)
L-M Tone	−0.1223*** (0.0374)	−0.1223*** (0.0399)	−0.1123*** (0.0340)	−0.1154*** (0.0206)
SASB Density	0.0034 (0.0089)	0.0034 (0.0051)	0.0085 (0.0073)	0.0087* (0.0048)
SASB HHI Balance	0.0049 (0.0465)	0.0049 (0.0335)	0.0337 (0.0485)	0.0203 (0.0310)
Log Market Cap	−0.0068 (0.0088)	−0.0068** (0.0029)	−0.0118* (0.0071)	−0.0066* (0.0036)
Integrated Report	0.0330** (0.0164)	0.0330*** (0.0088)	0.0284 (0.0181)	0.0485*** (0.0109)
Log Word Count	−0.0253*** (0.0084)	−0.0253*** (0.0089)	−0.0238*** (0.0088)	−0.0234*** (0.0051)
Observations	390	390	390	390
AR(1) ρ	—	—	0.4500	—

*Notes: Estimates of the preferred Model 1c specification under four alternative error structures. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Column 1 reproduces the baseline pooled OLS with bank-clustered standard errors. Column 2 reports Beck-Katz Panel-Corrected Standard Errors, which allow panel-specific heteroskedasticity and contemporaneous cross-bank correlation in the residuals while retaining OLS point estimates. Column 3 reports Feasible GLS with a common AR(1) coefficient on the within-bank residual process (Prais-Winsten transformation; estimated $\rho = 0.4500$). Column 4 reports Feasible GLS with panel-specific heteroskedasticity only. Year and region fixed effects are included in all specifications. Sample: 80 listed commercial banks, 2020–2024 ($N = 390$). The L-M Tone effect is significant at the 1% level under all four estimators; the Fog Index effect is significant at the 1% level under PCSE and panel-heteroskedasticity FGLS, at the 5% level under OLS with clustered SE, and at the 10% threshold under the FGLS-AR(1) specification ($p = 0.088$) — significant under all four estimators. The substantive null on SASB Density and SASB HHI Balance is broadly robust to all four estimators.*

Exclusion of 2020 (R1). When the 2020 observations are dropped — removing the year of greatest exogenous volatility — the L-M Tone effect remains significant at the 5% level ($\beta = -0.098$, $p = 0.012$) and the Fog Index effect strengthens to significance at the 5% level ($\beta = +0.0077$, $p = 0.020$) on the restricted sample ($N = 314$). The model R^2 declines substantially from 0.53 to 0.21, confirming that the 2020 year dummy carries a large share of the model's explanatory power, but the strengthening of both stylistic effects on the

COVID-excluded sample provides reassurance that neither result is an artefact of the COVID year.

Alternative readability measure (R2). Substituting the Flesch-Kincaid Grade level for the Fog Index leaves the L-M Tone effect essentially unchanged ($\beta = -0.119$, $p = 0.002$), while the Flesch-Kincaid coefficient is marginally significant at the 10% level ($\beta = +0.0049$, $p = 0.092$). The point estimate is positive — directionally consistent with the obfuscation channel — but smaller in magnitude than the Fog Index coefficient. This suggests that the Fog Index, which weights syllable complexity more heavily than the Flesch-Kincaid Grade, is the preferred readability proxy for this sample. The robustness of the tone effect to readability specification is reassuring for the main result.

Standalone reports only (R3). Restricting the sample to standalone ESG, CSR and sustainability reports ($N = 285$) strengthens both core results substantially: the Tone effect strengthens to $\beta = -0.203$ ($p < 0.001$) and the Fog effect strengthens to $\beta = +0.0068$ ($p = 0.029$). This is one of the most interpretively important robustness results in the analysis: both stylistic effects on volatility are approximately 70% larger in dedicated ESG reports than in the pooled sample. Standalone ESG reports are produced under a different editorial logic from integrated reports: they are explicitly framed as sustainability documents, are reviewed by sustainability functions rather than investor-relations functions, and are typically targeted at a different audience. The amplification of both effects in this subsample is consistent with the signalling interpretation of H1: when banks produce a document whose primary purpose is sustainability communication, the textual style they adopt carries a more direct signal about the underlying ESG performance than the same stylistic choices embedded in a financially focused annual report.

Table 8: Selective Obfuscation — Fog and Tone in Material vs Non-Material Text

Within-document paired comparison across 390 bank-year reports

Panel A: Group means and gap statistics

	Material text	Non-material text	Within-document gap
Fog Index — mean (SD)	30.11 (6.09)	20.57 (1.62)	+9.54 (5.41)
L-M Tone — mean (SD)	+0.098 (0.26)	+0.370 (0.21)	-0.272 (0.19)
Material content share	12% of sentences / 19% of words	—	—

Panel B: Tests of the within-document gap (H_0 : mean gap = 0)

Test	Fog gap	Tone gap
Paired t-test (t, p)	$t = +34.8$, $p < 10^{-120}$	$t = -28.3$, $p < 10^{-95}$
Wilcoxon signed-rank (W, p)	$W = 3$, $p < 10^{-65}$	$W = 2538$, $p < 10^{-57}$

Bank-clustered one-sample (t, p)	t = +20.4, p < 10 ⁻⁹²	t = -16.2, p < 10 ⁻⁵⁸
Cohen's d (paired effect size)	+1.76 (very large)	-1.43 (very large)
Direction agreement with mean	388 / 390 (99.5%)	356 / 390 (91.3%)

Panel C: Sub-sample breakdown (bank-clustered one-sample test on gap)

Sub-sample	N	Fog gap	Tone gap
Year 2020	76	+9.96***	-0.32***
Year 2021	79	+9.94***	-0.29***
Year 2022	80	+10.23***	-0.27***
Year 2023	78	+8.64***	-0.25***
Year 2024	77	+8.91***	-0.23***
Standalone reports	285	+9.65***	-0.26***
Integrated reports	105	+9.24***	-0.30***

*Notes: Within each report, sentences containing at least one SASB primary-keyword match are classified as 'material'; the remainder are 'non-material'. Fog Index and L-M Tone are then computed separately on each subset using the same algorithms applied to the full document, and the within-document gap is the material minus non-material value. *** p<0.01, ** p<0.05, * p<0.10. The pattern of within-document obfuscation and de-tonalisation of material content is present in 99.5% of reports (Fog) and 91.1% of reports (Tone), is highly significant under every test, exhibits a Cohen's d well above 1.0 in magnitude, and is consistent across all five years and across both standalone and integrated report types. The pattern is documentary of a systematic disclosure behaviour rather than an artefact of a particular bank, year or document format.*

Regional subsample (R4). Restricting the sample to Asia-Pacific banks yields a significant tone effect at the 5% level. The Fog effect remains insignificant in this subsample, and the SASB metrics are insignificant as expected.

Bank size control (R5). With log total assets included as an explicit size control alongside log market capitalisation, the Tone effect strengthens to $\beta = -0.130$ (p = 0.001) and the Fog effect strengthens to the 10% threshold ($\beta = +0.0064$, p = 0.056). Log total assets is small and not significant ($\beta = +0.010$, p = 0.412), a reasonable finding given that market capitalisation also captures bank size and the two measures are correlated. The strengthening of both the Tone and the Fog coefficient on the size-controlled sample, together with the corroboration from the impression-management evidence on the readability channel, provides convergent evidence that the H1 findings are not artefacts of an omitted bank-size variable.

Table 9: Topic-Conditioned Disclosure Quality in the Volatility Model

Dependent Variable: Annualised Stock Return Volatility

Variable	1c baseline	A: replace	B: add gaps	C: horse race
Fog Index (whole document)	0.0072**		0.0081**	0.0088*
	(0.0031)		(0.0036)	(0.0045)
Fog Index — material text		0.0012		-0.0009
		(0.0010)		(0.0014)
Fog gap (material – non-material)			-0.0008	

			(0.0013)	
L-M Tone (whole document)	−0.1223***		−0.1265***	−0.1524**
	(0.0374)		(0.0379)	(0.0646)
L-M Tone — material text		−0.0487*		0.0257
		(0.0278)		(0.0475)
Tone gap (material – non-material)			0.0191	
			(0.0386)	
SASB Density	0.0034	0.0136	0.0012	0.0003
	(0.0089)	(0.0098)	(0.0096)	(0.0099)
SASB HHI Balance	0.0049	−0.0292	0.0107	0.0101
	(0.0465)	(0.0472)	(0.0510)	(0.0504)
Log Market Cap	−0.0068	−0.0085	−0.0077	−0.0077
	(0.0088)	(0.0090)	(0.0090)	(0.0090)
Integrated Report	0.0330**	0.0347**	0.0328*	0.0325*
	(0.0164)	(0.0156)	(0.0171)	(0.0171)
Log Word Count	−0.0253***	−0.0136*	−0.0261***	−0.0259***
	(0.0084)	(0.0082)	(0.0081)	(0.0081)
Year Fixed Effects	Yes	Yes	Yes	Yes
Region Fixed Effects	Yes	Yes	Yes	Yes
Observations	390	390	390	390
R-squared	0.5305	0.5118	0.5319	0.5320
Adjusted R-squared	0.5117	0.4922	0.5105	0.5106
AIC	−854.0	−838.8	−851.1	−851.2

Notes: Standard errors clustered at the bank level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Column 1 reproduces Model 1c. Column 2 ('A: replace') substitutes the Fog and L-M Tone of material sentences for the whole-document equivalents. Column 3 ('B: add gaps') retains the whole-document measures and adds the within-document gap variables. Column 4 ('C: horse race') includes both the whole-document and the material-text measures jointly. Among the topic-conditioned variables, only tone_material is marginally significant at 10% in Spec A; the gap measures fog_gap and tone_gap are insignificant in Spec B, and neither material-text measure is significant in the horse-race Spec C. The whole-document Fog and L-M Tone effects survive their re-estimation in columns 3 and 4. Together with the descriptive evidence in Table 8, this indicates that selective obfuscation and de-tonalisation of material content is a structural feature of bank ESG reporting (present in 99.5% and 91.3% of reports respectively) but is not informative for cross-bank variation in stock-return volatility — the disclosure-quality signal investors appear to price lives in the whole-document distribution, not in the within-document material-vs-non-material gap. Sample: 80 listed commercial banks, 2020–2024 ($N = 390$).

Estimator choice. Mundlak decomposition and a regression-based Hausman test. A natural question is whether the disclosure-quality results survive bank fixed effects. A within (fixed-effects) estimator is, however, poorly suited to this setting: the disclosure-quality measures are highly persistent within a bank over the five-year window, so bank fixed effects discard precisely the cross-sectional variation that identifies the hypothesis and retain only noisy year-to-year fluctuations. To settle the estimator choice formally rather than by assertion, the full specification is re-estimated as a Mundlak (1978) correlated-random-effects model, augmenting the pooled regression with the bank-level means of every time-varying regressor (Table 10). In this framework the coefficient on each variable recovers the within-bank (fixed-effects-equivalent) effect, the coefficient on its bank mean captures the

contextual between-bank effect, the between effect is the sum of the two, and a joint Wald test that all of the mean terms are zero is a cluster-robust analogue of the Hausman test. The omnibus test rejects the null that the bank means are jointly zero ($F = 3.47$, $p = 0.001$), but the rejection is driven almost entirely by return on equity, whose within-bank effect (-0.43) and between-bank effect (+0.08) diverge sharply because profitability is strongly cyclical within banks. When the test is restricted to the four disclosure-quality variables it no longer rejects ($F = 1.97$, $p = 0.107$): the within and between disclosure effects are statistically indistinguishable, so pooled OLS — which exploits both sources of variation — yields consistent estimates of the disclosure coefficients and bank fixed effects are not required to identify them. The decomposition further shows that the Fog Index effect is essentially a between-bank phenomenon (within coefficient approximately zero; contextual coefficient +0.0081, $p < 0.05$), confirming that disclosure quality is a persistent institutional characteristic for which the within estimator is uninformative over so short a panel. Consistent with this, the standalone bank-fixed-effects estimates render both the Tone and Fog coefficients insignificant while a random-effects estimator returns an intermediate, marginally significant tone coefficient (beta = -0.073, $p = 0.069$); the F-test for poolability rejects pooling ($F = 4.61$, $p < 0.001$), confirming that bank-level heterogeneity is present but, as the focused Hausman test shows, not correlated with the disclosure regressors in a way that biases the pooled estimates. One within-bank result does emerge and is retained from the fixed-effects analysis: SASB Density carries a positive and significant within coefficient (+0.0165, $p < 0.01$ in the Mundlak decomposition; +0.029 under standard bank fixed effects), indicating that in years when a bank expands its material-topic disclosure its volatility is higher. This most plausibly reflects reverse causality — banks broaden risk and materiality disclosure in more turbulent years — and is treated as a boundary on the otherwise cross-sectional interpretation of the results.

Table 10: Mundlak Decomposition and Regression-Based Hausman Test

Dependent Variable: Annualised Stock Return Volatility

Variable	Within (FE)	Contextual (mean)	Between
Fog Index	-0.0020 (0.0025)	0.0081** (0.0036)	0.0060
L-M Tone	-0.0388 (0.0265)	-0.0451 (0.0474)	-0.0840
SASB Density	0.0165*** (0.0059)	-0.0175* (0.0106)	-0.0010
SASB HHI Balance	0.0686**	-0.0415	0.0272

	(0.0345)	(0.0633)	
Log Market Cap	0.0031	-0.0190	-0.0159
	(0.0135)	(0.0147)	
ROE	-0.4256***	0.5058**	0.0802
	(0.1115)	(0.2084)	
Beta	0.2201***	0.0233	0.2434
	(0.0134)	(0.0291)	
Tier 1 Capital Ratio	-0.0340	-0.2354	-0.2695
	(0.3002)	(0.3406)	
Log Word Count	0.0051	-0.0273***	-0.0222
	(0.0063)	(0.0098)	

*Notes: Pooled OLS with bank-clustered standard errors (in parentheses), augmented with the bank-level (entity) means of every time-varying regressor (Mundlak, 1978; correlated random effects). The coefficient on each variable is the within-bank, fixed-effects-equivalent estimate; the coefficient on its bank mean is the contextual effect; the between-bank effect equals their sum (reported in the final column). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Year and region fixed effects and a constant are included; the bank-mean coefficients are estimated but only their joint significance is reported below. A joint Wald test that all nine bank-mean coefficients equal zero — the regression-based Hausman test — rejects ($F = 3.47$, $p = 0.001$), but the rejection is driven by return on equity, whose within (-0.4256) and between (0.0802) effects diverge sharply because profitability is strongly cyclical within banks. Restricting the test to the four disclosure-quality means does not reject ($F = 1.97$, $p = 0.107$), indicating that the within and between disclosure effects are statistically indistinguishable and that pooled OLS yields consistent estimates of the disclosure-quality coefficients. The Fog Index effect operates entirely between banks (within ≈ 0 ; contextual = 0.0081 , $p < 0.05$), confirming that disclosure quality is a persistent bank characteristic for which bank fixed effects — which discard all between-bank variation — are uninformative over the five-year window. Sample: 80 listed commercial banks, 2020–2024 ($N = 390$; $R^2 = 0.8012$).*

Temporal ordering and reverse causality: one-year lags. To probe the temporal ordering of the disclosure-volatility relationship, the model is re-estimated with the six report-derived variables (Fog, Tone, SASB Density, SASB HHI Balance, the integrated-report dummy and log word count) replaced by their one-year lags — the metrics of each bank's prior-year report — while the financial controls remain contemporaneous (Table 11). The lag necessarily drops the 2020 reporting wave, which has no prior-year report, and one bank (Barclays, whose 2020/2022/2024 coverage admits no consecutive one-year lag), leaving 307 bank-year observations; the contemporaneous and lagged columns of Table 11 are estimated on this identical sample, so that only the timing of the disclosure measures differs. The Fog Index coefficient is virtually unchanged under lagging ($+0.0044$ lagged versus $+0.0045$ contemporaneous) and retains its significance, indicating that prior-year report obfuscation predicts current-year volatility; because current volatility cannot have caused a report written a year earlier, this materially strengthens the causal interpretation of the readability channel and mitigates the reverse-causality concern for H1a. The L-M Tone coefficient, by contrast, attenuates by roughly 40% and loses statistical significance when lagged (-0.027 lagged versus -0.046 contemporaneous), indicating that the tone-volatility link is primarily contemporaneous and should be read as a concurrent association rather than as evidence that report tone forecasts future risk. The SASB measures remain insignificant

under both timings. The contrast is itself informative: the readability channel has a forward-looking, reverse-causality-resistant component, whereas the tone channel is concurrent — a distinction carried into the identification discussion in Section 6.2.

Alternative error-structure estimators (Table 7). The preferred Model 1c specification is re-estimated under three alternative residual error structures. Beck-Katz Panel-Corrected Standard Errors (Beck & Katz, 1995) allow panel-specific heteroskedasticity and contemporaneous cross-bank correlation in the residuals while retaining the OLS point estimates; this estimator is the canonical choice for time-series cross-section panels in which year fixed effects do not exhaust the common time-shock structure. Feasible GLS with a common AR(1) coefficient on the within-bank residual process — implemented via the Prais-Winsten transformation, with an estimated ρ of 0.45 — generates new and more efficient coefficient estimates if the AR(1) error structure is correctly specified. Feasible GLS with panel-specific heteroskedasticity only generates weighted-least-squares estimates with weights equal to the inverse of each bank's residual variance. The results across all four estimators tell a consistent story for the headline findings. The L-M Tone coefficient is significant at the 1% level under every estimator (β ranging from -0.112 to -0.122), confirming the robustness of the H1b finding to alternative assumptions about the residual covariance structure. The Fog Index coefficient is significant at the 1% level under Beck-Katz PCSE and FGLS with panel heteroskedasticity, at the 5% level under OLS with clustered standard errors, and at the 10% threshold under FGLS with AR(1) errors ($\beta = +0.0050$, $p = 0.088$); the AR(1) point estimate is consistent with persistent within-bank readability differences that the AR(1) transformation differences out. The substantive null on SASB Density and SASB HHI Balance is robust to all four estimators.

Topic-conditioned disclosure quality (Tables 8 and 9). The descriptive evidence reported in Section 5.1 — that material content accounts for only 12% of sentences and 19% of words across the sample, and that the Fog Index of material sentences is on average 9.5 points higher and the L-M Tone of material sentences is on average 0.27 lower than the corresponding measures for non-material content — motivates a topic-conditioned reformulation of the H1 specification. Specifically, `fog_material` and `tone_material` are constructed by computing the Fog Index and L-M Tone on the subset of each report consisting only of sentences containing at least one SASB primary keyword; their non-material complements `fog_non_material` and `tone_non_material` are computed on the

remaining sentences; and the within-document gap measures `fog_gap` and `tone_gap` are defined as the corresponding differences. Three alternative specifications are reported in Table 9 alongside the Model 1c baseline: Spec A replaces the whole-document Fog and Tone with their material-text equivalents; Spec B retains the whole-document measures and adds the gap variables; and Spec C includes both the whole-document and the material-text measures jointly (a "horse race" specification). The descriptive evidence on the size and universality of the within-document gaps is overwhelming: paired t-tests on the Fog gap and the Tone gap reach t-statistics of +34.8 and −28.3 respectively, with p-values below 10^{-95} ; the corresponding bank-clustered tests yield t-statistics of +20.4 and −16.2, p-values below 10^{-58} ; Cohen's d effect sizes are +1.76 and −1.43, both in the very-large category; the Fog gap is positive in 99.5% of reports and the Tone gap is negative in 91.3%; and the pattern is present in every year of the sample and in both standalone and integrated reports. Yet the regression evidence in Table 9 indicates that none of the topic-conditioned variables — `fog_material`, `tone_material`, `fog_gap`, `tone_gap` — is significant at the 5% level in any specification (`tone_material` is marginally significant at 10% in Spec A only). The whole-document Fog and Tone effects survive their re-estimation in Specs B and C (whole-document Tone $\beta = -0.127$ in Spec B, $p < 0.01$; whole-document Fog $\beta = +0.0081$, $p < 0.05$). The combined evidence indicates that selective obfuscation and de-tonalisation of material content is a structural feature of bank ESG reporting present in almost every report but does not provide additional cross-sectional information about which banks investors should consider more risky. The disclosure-quality signal that the cross-section appears to price lives in the whole-document distribution of Fog and Tone, not in differential treatment of material content within each report — even though that differential treatment is itself a striking and previously undocumented descriptive feature of the data. The substantive interpretation of this combined evidence is taken up in Sections 6.1 and 6.3.

Table 11: One-Year-Lag Robustness — Prior-Year Disclosure Quality and Volatility

Dependent Variable: Annualised Stock Return Volatility (year t)

Variable	Contemporaneous (t)	Lagged (t − 1)
Fog Index	0.0045*	0.0044*
	(0.0023)	(0.0025)
L-M Tone	−0.0462*	−0.0270
	(0.0280)	(0.0275)
SASB Density	0.0001	−0.0005
	(0.0066)	(0.0069)
SASB HHI Balance	0.0238	0.0055

	(0.0328)	(0.0349)
Observations	307	307
R-squared	0.6311	0.6282

*Notes: Standard errors clustered at the bank level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. In the 'Lagged ($t - 1$)' column the six report-derived variables (Fog Index, L-M Tone, SASB Density, SASB HHI Balance, the integrated-report dummy and log word count) are replaced by their one-year lags — the metrics of the bank's prior-year report — while the financial controls (log market capitalisation, ROE, beta and the Tier 1 capital ratio) remain contemporaneous. Both columns are estimated on the identical lagged sample ($N = 307$): the 2020 reporting wave has no prior-year report, and Barclays drops out because its 2020/2022/2024 coverage admits no consecutive one-year lag. The only difference between the columns is therefore the timing of the disclosure measures. Year fixed effects (rebased to 2021) and region fixed effects are included. The Fog Index coefficient is essentially identical under lagging and retains its significance, indicating that prior-year report obfuscation predicts current-year volatility and mitigating reverse-causality concerns; the L-M Tone coefficient attenuates by roughly 40% and loses significance, indicating that the tone–volatility association is primarily contemporaneous.*

5.6 Summary of Empirical Results

The empirical evidence presented in this chapter can be summarised in the following statements.

The linguistic tone of bank ESG disclosures — measured by the Loughran-McDonald net positive sentiment ratio — is robustly and negatively associated with annualised stock return volatility in the cross-section. The tone effect is significant at the 1% level in Model 1c ($\beta = -0.122$, $p = 0.001$) and at the 5% level once year-varying ROE and beta are added (Model 1b, $\beta = -0.055$, $p = 0.030$). It strengthens further under explicit bank-size control (R5, $\beta = -0.130$, $p = 0.001$), is amplified in dedicated standalone ESG reports (R3, $\beta = -0.203$, $p < 0.001$), survives exclusion of the high-volatility COVID year (R1, $\beta = -0.098$, $p = 0.012$), is significant at the 5% level in the Asia-Pacific subsample (R4), and is significant at the 1% level under all four alternative error-structure estimators examined. Hypothesis H1b is strongly supported.

The readability of bank ESG disclosures — measured by the Gunning Fog Index — exhibits a positive relationship with stock return volatility consistent with the obfuscation channel of H1a. The Fog effect is significant at the 5% level in Model 1c ($\beta = +0.0072$, $p = 0.022$), in the univariate Model 1d ($\beta = +0.0067$, $p = 0.048$) and in the standalone-reports subsample (R3, $\beta = +0.0068$, $p = 0.029$), at the 10% level in the full-control Models 1a/1b that include year-varying beta and ROE and in the bank-size-controlled robustness check R5, and at the 10% level or better under all four alternative error-structure estimators. It is corroborated by a strongly significant impression-management result that banks with lower ROE produce

systematically less readable reports (H2a, $\beta = -5.55$, $p = 0.028$). Hypothesis H1a is supported.

Log market capitalisation, denominated in year-end USD, enters Models 1a and 1b with the theoretically expected negative sign and is highly significant ($\beta \approx -0.015$, $p \approx 0.007$), confirming the standard size–risk relationship in banking. Year-varying ROE enters with a strongly negative coefficient and is significant at the 1% level in both Models 1a and 1b ($\beta \approx -0.21$, $p \approx 0.008$).

When year-varying market beta is added as a control in Models 1a and 1b, the tone coefficient remains significant at the 5% level ($\beta = -0.055$, $p = 0.030$ in Model 1b), and the Fog coefficient retains marginal-10% significance ($\beta = +0.0034$, $p = 0.078$ in Model 1b). The retention of both stylistic effects after partialling out year-matched systematic risk and operating performance indicates that disclosure quality operates as an independent channel for investor uncertainty rather than as a passive correlate of financial fundamentals.

The disclosure-quality findings are robust to the inclusion of the Tier 1 capital ratio as a solvency control in the augmented full specification (Model 1b+T1). The capital ratio is correctly signed but insignificant (beta = -0.194, $p = 0.246$), and the Fog and Tone coefficients are essentially unchanged by its inclusion (Fog +0.0033, $p = 0.085$; Tone -0.0566, $p = 0.024$).

The choice of pooled OLS over bank fixed effects is supported by a formal test. A Mundlak correlated-random-effects decomposition shows that the disclosure effects load on the between-bank (contextual) component, and a regression-based Hausman test restricted to the four disclosure variables fails to reject the consistency of the pooled estimator ($F = 1.97$, $p = 0.107$). The omnibus Hausman rejection ($F = 3.47$, $p = 0.001$) is attributable to the cyclical within-bank behaviour of ROE, not to the disclosure regressors. Bank fixed effects are therefore not required to identify the cross-sectional disclosure effects, which is the central interest of the thesis.

The readability and tone channels differ in their temporal ordering. Under a one-year lag of the report-derived variables, the Fog Index effect is essentially unchanged and remains significant (+0.0044 lagged versus +0.0045 contemporaneous), strengthening the causal reading of the obfuscation channel, whereas the L-M Tone effect attenuates and loses

significance (-0.027 lagged versus -0.046 contemporaneous), indicating that the tone-volatility association is contemporaneous.

The tone-volatility relationship is contingent on bank performance. A Tone x ROE interaction is positive and significant ($\beta = 0.717$, $p = 0.039$): tone is strongly associated with lower volatility for weak performers (-0.086 at one standard deviation below mean ROE, $p = 0.007$) but is insignificant for strong performers (-0.023, $p = 0.353$). The tone effect is thus concentrated among the banks with the greatest incentive to manage impressions, complementing the direct H2 evidence.

The impression management hypothesis is supported through the readability channel and not supported through the tone channel. The obfuscation channel — financially weaker banks producing less readable reports — is supported at the 5% significance level (ROE $\rightarrow$ Fog $\beta = -5.55$, $p = 0.028$). The tone-inflation channel is rejected: the coefficient on ROE in the tone regression is positive but insignificant ($\beta = +0.366$, $p = 0.175$). The substantive content of SASB disclosure is not significantly related to ROE in either direction. The two stylistic dimensions of impression management thus operate independently rather than as twin manifestations of a single managerial impulse.

The substantive content metrics — ESG vocabulary density, SASB material topic density, SASB HHI Balance and SASB Proportional Coverage — are not significantly associated with stock return volatility in the cross-sectional specifications. The materiality premium hypothesis (H3) is not supported: ESG and SASB density measures yield essentially identical model fit ($\Delta R^2 < 0.0001$), reflecting both their high collinearity ($r = 0.854$) and the apparent indifference of investors to the substantive-content dimension of disclosure as currently measured. The within-bank fixed-effects specification, however, surfaces a positive and significant SASB-density coefficient ($\beta = +0.029$, $p \approx 0.012$; $\beta = +0.018$, $p \approx 0.032$ under random effects), suggesting that changes in material-topic disclosure across time within a bank are associated with elevated idiosyncratic volatility — a pattern worth following up in future research with a longer panel.

A topic-conditioned analysis of the textual material documents a near-universal pattern of selective obfuscation and de-tonalisation of material content. In 99.5% of reports the Fog Index of material sentences exceeds the Fog Index of the rest of the report (mean gap +9.5, Cohen's $d = +1.76$); in 91.3% of reports the L-M Tone of material content is less positive

than the tone of non-material content (mean gap -0.27 , Cohen's $d = -1.43$). The pattern is significant under every test applied (paired t , Wilcoxon signed-rank, bank-clustered one-sample tests) and consistent across years and report types. Yet cross-bank variation in this within-document gap does not independently predict stock-return volatility once the whole-document Fog and Tone measures are included: the gap is a level feature of the industry rather than a discriminator across banks. This is the strongest descriptive finding of the thesis, and it provides direct policy-relevant evidence that banks systematically engage in selective obfuscation of material content — even if cross-sectional variation in that practice is not itself priced by markets.

6. Discussion & Conclusion

6.1 Summary of Findings and Theoretical Implications

This thesis set out to test whether the quality of bank ESG disclosures — operationalised through four theoretically motivated text-analytic metrics — reduces stock return volatility. Across 390 bank-year observations spanning 80 banks and five years, the analysis yields a clear empirical picture. Linguistic credibility, captured by the Loughran-McDonald tone of net positive sentiment, is robustly and negatively associated with volatility in the cross-section, with the effect surviving conventional sample- and specification-robustness checks, alternative error-structure estimators, and the inclusion of year-varying beta and ROE. Linguistic accessibility, captured by the Gunning Fog Index, exhibits a complementary pattern: less readable reports are associated with higher volatility, and this readability effect is corroborated by the strong impression-management result that lower-ROE banks produce systematically less readable reports. The substantive content of disclosure, captured by ESG and SASB density measures, does not independently predict volatility in the cross-section. These cross-sectional disclosure-quality results are robust to the progressive enrichment of the financial-control set, including the addition of a year-varying Tier 1 capital ratio as an explicit solvency control, and the choice of pooled OLS over bank fixed effects is supported by a formal Mundlak/Hausman test rather than asserted. Two further refinements sharpen the interpretation of the tone finding: under a one-year lag the tone effect is contemporaneous (whereas the readability effect is forward-looking), and the tone effect is performance-contingent, concentrated among weaker performers.

The classical impression management hypothesis is partially supported. Banks with lower ROE produce systematically less readable reports, consistent with the obfuscation channel; but the data do not support the tone-inflation channel — no significant relationship between performance and tone is detected in either direction. The two stylistic dimensions of impression management thus operate independently. The strong cross-sectional tone–volatility relationship documented for H1b is therefore not the result of a performance-driven tone-inflation strategy but reflects either a credible-signalling channel — in which banks adopt tone consistent with their underlying disclosure quality — or a residual cross-sectional association whose precise causal interpretation remains an open question for future research.

A separate descriptive contribution emerges from the topic-conditioned analysis of disclosure quality. Bank ESG reports systematically and very strongly obfuscate and de-tonalise their material content: in 99.5% of the 390 reports examined, the Fog Index of sentences referencing SASB material topics exceeds the Fog Index of the rest of the report by an average of 9.5 points; in 91.3% of reports, the L-M Tone of material content is less positive than the tone of non-material content by an average of 0.27. Both effects are very large by conventional standards (Cohen's $d = +1.76$ and -1.43 respectively), highly significant under every test applied, and consistent across years and report types. This finding documents, for the first time in a global banking sample, a systematic disclosure behaviour that has not been previously quantified. The associated regression evidence indicates that the cross-bank variation in the within-document gap is small relative to its level: the obfuscation of material content is a near-universal industry practice, not a differentiator across banks. The cross-sectional disclosure-quality signal that investors appear to price therefore lives in the whole-document distribution of Fog and Tone, not in differential treatment of material content — even though the latter is itself a striking and previously undocumented feature of the data.

These findings carry three theoretical implications for the disclosure-quality literature. First, they reinforce the distinction between disclosure quantity and disclosure quality emphasised in the recent literature (Drempetic, Klein & Zwergel, 2020) and demonstrate that quality, when carefully operationalised through multiple text-analytic dimensions, captures variation that simple disclosure indicators or third-party ESG scores cannot. The two stylistic dimensions — readability and tone — predict volatility independently of the substantive-content dimensions in this sample, suggesting that investors attend to how banks communicate at least as much as to what they communicate. Second, the asymmetric support for the impression management hypothesis — strong in the readability channel, absent in the tone channel — implies that managers do not use the full stylistic repertoire of impression management uniformly when performance deteriorates. Obfuscation through complexity is a distinct strategic behaviour from positive-tone inflation, and the two should not be treated as twin manifestations of a single managerial impulse. Third, the failure of the materiality premium hypothesis (H3) in the banking sector tempers the optimism associated with the Khan, Serafeim and Yoon (2016) findings: while material performance may be priced by markets, this study finds no equivalent premium for material disclosure

once disclosure quality is operationalised through textual coverage. The selective-obfuscation evidence documented in Tables 8 and 9 helps interpret this null: investors may rationally place limited weight on the textual coverage of material topics if the underlying narrative is itself systematically obscured and de-tonalised relative to the surrounding text.

6.2 Interpretation of the Tone Finding

The negative association between L-M Tone and stock return volatility is the most robust empirical finding in this thesis, surviving every robustness check that retains adequate cross-sectional identifying variation and significant at the 1% level under all four alternative error-structure estimators examined. Its theoretical interpretation, however, is not immediately obvious because three distinct mechanisms could in principle produce the observed sign.

The first interpretation — credible signalling — is the one most consistent with the joint pattern of evidence. Under this account, banks with genuinely strong ESG and operating performance adopt unambiguously positive language in their disclosures, while banks with weaker performance retreat to hedged or cautious tone. Investors, in turn, interpret unambiguous positive language as a credible signal of underlying strength, reducing the dispersion of their assessments and dampening return volatility. Two pieces of evidence are consistent with this interpretation. First, the tone effect is approximately 70% stronger in standalone ESG reports ($R3, \beta = -0.203, p < 0.001$) than in the pooled sample (Model 1c, $\beta = -0.122$), consistent with the signal being most informative in documents whose primary purpose is sustainability communication. Second, the absolute-tone regression (Section 5.3) finds no significant relationship between $|\text{Tone}|$ and ROE, indicating that the relationship is not driven by extreme observations in either tail. The direct H2b test of tone inflation also fails to find evidence of strategic positive framing by weak performers ($\beta = +0.366, p = 0.175$), removing one obvious alternative to the credible-signalling account.

The second interpretation — credibility loss from over-optimism — would predict a non-monotonic relationship in which extremely high positive tone is discounted by investors and produces higher volatility than moderately positive tone. The linear specification in this thesis cannot directly test for non-monotonicity, and a quadratic-tone specification is identified as an avenue for future work. However, the data do not support the conjecture that

uniform positivity is discounted: tone is monotonically associated with volatility across the empirical range observed (-0.13 to $+0.69$), with no apparent inflection point in residual diagnostics.

The third interpretation — reverse causation through fundamentals — holds that banks with low fundamental volatility (driven by external factors such as state ownership, regulatory backing or business-mix stability) simply have fewer negative events to report, producing more positive prose as a passive consequence. This mechanism cannot be fully ruled out without an exogenous instrument for tone and represents the principal identification threat to the H1b finding. Three considerations partially mitigate the concern. First, the inclusion of beta — a forward-looking market measure of systematic risk — as a control absorbs much of the structural cross-bank variation in volatility. Second, year-varying ROE absorbs much of the operating-performance variation that would generate the reverse-causal mechanism. Third, the residual cross-sectional association between tone and volatility, after these controls, is most naturally interpreted as the marginal effect of disclosure-style variation that is independent of fundamentals. The one-year-lag analysis in Section 5.5 bears directly on this identification question and yields a split verdict across the two stylistic channels. The Fog Index effect is essentially identical when prior-year readability is used to explain current-year volatility, which — because current volatility cannot cause a report written a year earlier — substantially strengthens the causal reading of the obfuscation channel and is corroborated by the impression-management evidence that lower-ROE banks produce less readable reports. The tone effect, by contrast, attenuates and loses significance under lagging, indicating that the tone-volatility relationship is contemporaneous. This does not establish reverse causation for tone, but it does mean the tone finding should be read as a contemporaneous cross-sectional association rather than as evidence that tone forecasts future risk, and it raises the evidentiary bar for the credible-signalling interpretation relative to the readability result. Nonetheless, full causal identification would require an external instrument or a natural experiment — an issue addressed in Section 6.4.

6.3 Interpretation of the Asymmetric Impression Management Pattern

The finding that obfuscation operates as predicted while tone inflation does not has implications beyond the specific hypotheses of this thesis. It suggests that the impression-management toolkit available to managers is not used uniformly across stylistic dimensions, and that readability and tone carry distinct strategic logics. The point is reinforced by the descriptive selective-obfuscation evidence in Sections 5.1 and 5.5: managers systematically obfuscate material content at the within-document level (Fog gap +9.5, present in 99.5% of reports) but do not systematically de-tonalise material content to a smaller degree when their financial performance is weak.

Obfuscation through complexity has a strong defensive logic. When financial performance is weak, managers gain by slowing the rate at which adverse information is processed by markets, analysts and civil society organisations. Higher-syllable-density, longer-sentence prose imposes processing costs on readers, plausibly muting the immediate market reaction to bad news (Bloomfield, 2002). The observed coefficient ($\text{ROE} \rightarrow \text{Fog } \beta = -5.55, p = 0.028$) implies that banks with the lowest ROE in the sample produce ESG reports approximately 1.3 Fog units less readable than banks with the highest ROE — a meaningful difference equivalent to roughly two thirds of one cross-sectional standard deviation in readability. The within-document evidence from the topic-conditioned analysis is even more striking: regardless of overall performance, banks make their material-topic discussion approximately 9.5 Fog units harder to read than the surrounding text, a difference of approximately one cross-sectional standard deviation occurring within nearly every single report in the sample.

Positive-tone inflation, by contrast, has a more limited offensive logic. To work as impression management, positive tone must be credibly interpreted by markets as informative about underlying performance — but if markets are aware that weaker performers have incentives to inflate tone, the equilibrium response is to discount positive language from weak performers, removing the strategic benefit. The data are consistent with this discounting equilibrium: weaker-performing banks do not in fact produce more positive tone ($\beta = +0.366, p = 0.175$). Stronger performers, who do not face the same credibility-discount, can adopt unambiguous positive language and have it interpreted at face value — and the within-document tone gap pattern (material content less positive in 91.3% of reports,

regardless of overall performance) is consistent with markets having calibrated their expectations to this stylistic convention.

A further asymmetry concerns not whether managers inflate tone but where tone is priced. The Tone x ROE interaction (Section 5.3, Table 4B) shows that the volatility-dampening association of positive tone is concentrated among weak performers and vanishes for strong performers. This complements, rather than contradicts, the null on tone inflation. Weak performers do not on average adopt more positive tone; but among weak performers, those that do adopt positive tone are associated with markedly lower volatility, whereas for strong performers tone is uninformative. The pattern is consistent with tone functioning as a performance-contingent signal: where investor uncertainty about a bank is greatest — that is, where performance is weakest — the linguistic register of the ESG report has the most "work" to do in shaping the dispersion of investor beliefs, while for strong performers the fundamentals speak for themselves and the report's tone adds little. Read alongside the obfuscation result, this paints a coherent picture of how the two stylistic channels operate under weak performance: managers of weak banks systematically reduce readability (the obfuscation channel), and the tone they adopt is precisely where the market's volatility response to tone is concentrated.

The policy and regulatory implications of this asymmetry are taken up in Section 6.5. The core observation is that mandatory readability standards — analogous to the SEC's plain-English requirements for financial prospectuses — could constrain the obfuscation channel directly, and that a targeted standard requiring comparable readability for material-topic discussion would address a specific and systematically documented practice. Equivalent constraints on tone are neither feasible nor necessarily desirable.

6.4 Limitations

Five limitations of the analysis deserve specific acknowledgement, both as caveats to the interpretation offered above and as motivations for the future research directions identified in Section 6.6.

First, identification is fundamentally cross-sectional. The Mundlak decomposition and regression-based Hausman test reported in Section 5.5 establish this formally: the within-bank time-series variation in disclosure quality is too small, in a five-year panel of large established banks, to identify the cross-sectional effects of interest, and the disclosure effects load on the between-bank component. The focused Hausman test does not reject the consistency of the pooled estimator for the disclosure variables ($F = 1.97$, $p = 0.107$), so pooled OLS is appropriate; but the corollary is that the H1 findings are most appropriately interpreted as the cross-sectional association between disclosure quality and volatility rather than the causal effect of within-bank changes in disclosure quality. The one-year-lag analysis goes some way towards a temporal-ordering argument for the readability channel — prior-year Fog predicts current-year volatility — but the tone channel is contemporaneous, and full causal identification would still require an exogenous shock to disclosure quality (the staggered CSRD implementation in 2024-2025 being the leading candidate) or an instrument that affects disclosure quality without operating through volatility directly. Second, although all three primary financial controls (market capitalisation, beta and ROE) are now year-varying for the listed-bank sample, the panel covers only five years. A longer panel — and ideally a panel that exploits the staggered CSRD-implementation experiment in Europe (2024–2025) — would meaningfully improve causal identification, and the post-submission addition of 2025 report data already in preparation will deliver a first extension along this dimension.

Third, the analysis is limited to English-language report text. Reports published in languages other than English are excluded, and reports with substantial non-English content (notably for some Chinese banks where English versions are sometimes partial translations) may have their disclosure-quality metrics underestimated. A language-adjusted analysis using machine-translated content or native-language sentiment dictionaries is an avenue for future work.

Fourth, the SASB dictionary captures keyword presence and weighted density, but does not directly capture quantitative data disclosure. A bank that mentions "financed emissions" five times in its sustainability report contributes the same to the SASB Density metric as a bank that mentions the term once but reports detailed Scope 3 emissions figures in a structured data table. The null result on the materiality premium hypothesis may reflect this

measurement limitation rather than an underlying absence of an effect, and the development of a SASB metric that explicitly captures quantitative disclosure is a promising direction.

Fifth, the L-M dictionary, while purpose-built for financial texts, was developed and validated on annual reports rather than sustainability reports. The relative weight of positive and negative financial-sentiment words in ESG reports may differ from their weight in 10-Ks, and a sustainability-specific tone dictionary — currently under development by several research groups — would improve the precision of the tone measure. The fact that the tone effect is robust across multiple specifications using the L-M dictionary suggests this is not a critical concern, but it remains a refinement worth pursuing.

6.5 Policy and Regulatory Implications

The empirical findings have implications for three streams of ongoing regulatory activity in sustainability disclosure: the European Union's Corporate Sustainability Reporting Directive (CSRD) and the underlying European Sustainability Reporting Standards (ESRS); the International Sustainability Standards Board (ISSB) global baseline; and sector-specific guidance from banking regulators including the European Banking Authority's Pillar 3 ESG framework.

First, the convergent evidence for the obfuscation channel suggests that readability standards deserve serious consideration in mandatory sustainability disclosure frameworks. The SEC has long imposed plain-English requirements on Section 421(d) prospectuses and certain other investor communications, with the explicit goal of reducing information-processing costs for retail investors. Sustainability reports — increasingly used by institutional and retail investors alike, and frequently exceeding 200 pages — currently operate under no such constraint. A regulatory readability standard analogous to a maximum Fog Index threshold (e.g., requiring sustainability reports to score below 18 on the Fog scale, corresponding to undergraduate reading difficulty) would address the specific channel through which low-performing banks currently obscure adverse information from less-attentive readers. The selective-obfuscation evidence in Section 5.1 implies that an even more targeted regulatory leverage point would be a readability requirement applied specifically to material-topic discussion — for example, requiring that the subsection of

each sustainability report discussing financed emissions, business ethics or systemic risk management achieve no higher Fog Index than the report as a whole. Such a requirement would directly address the documented practice of selective obfuscation of material content, which is present in 99.5% of the reports in this sample and operates at a magnitude approximately one cross-sectional standard deviation in Fog.

Second, the failure of the SASB materiality premium has implications for the design of content-based disclosure frameworks. The CSRD and ESRS focus heavily on the substantive content of disclosure — what topics must be covered, what metrics must be reported — and have so far devoted comparatively less attention to the stylistic quality of the resulting reports. The empirical evidence suggests this prioritisation may be incomplete: investors do not currently appear to differentially price topic-level material coverage in the textual narrative, and additional coverage requirements may produce longer reports without commensurate informational value. The policy implication is not that content standards are misplaced, but that they should be paired with quality standards — readability standards, structured-data-format requirements, comparability standards — to translate coverage into actionable information.

Third, the credible-signalling interpretation of the tone effect implies that uniform positive framing — sometimes encouraged by communications functions seeking to project corporate confidence — may be counterproductive from an investor-uncertainty perspective. Banks that hedge their positive claims, acknowledge specific risks and adopt measured tone where the underlying performance is genuinely mixed are more likely to be rewarded by reduced volatility than those that adopt blanket positivity. This is a managerial implication rather than a regulatory one, but it has direct relevance for the editorial choices made by communications and sustainability functions in large banks.

6.6 Future Research

Six avenues for future research follow directly from this analysis.

First, the temporal coverage of the panel should be extended. The five-year window 2020-2024 constrains the within-bank identifying variation; extending the panel forward to 2025

(a step already in preparation) and backward to the 2017-2019 period would strengthen the bank fixed-effects estimates and enable cleaner identification of within-bank effects of disclosure-quality changes. The Tier 1 capital ratio, anticipated in earlier drafts as a leverage and capital-adequacy control, has been incorporated into the augmented specification (Model 1b+T1, Tables 3, 4B, 10 and 11); consistent with the tight cross-sectional clustering of capital ratios among large banks, it enters with the expected negative sign but is not statistically significant once systematic risk, profitability and size are controlled. A longer panel with greater within-bank variation in capital ratios — spanning, for example, the post-2025 Basel III "endgame" transition — would provide a more powerful test of the capital-adequacy channel than the current window allows.

Second, causal identification should be pursued through the natural experiment provided by the CSRD implementation in 2024–2025. European banks above specified size thresholds are required to publish CSRD-compliant sustainability reports for fiscal year 2024 onwards, mandating structured topic coverage, assured disclosure and (in principle) higher disclosure quality than under the previous voluntary regime. A difference-in-differences design comparing affected European banks to unaffected control-group banks (e.g., similar-sized non-EU banks) over the 2023–2025 window could provide causal evidence on the disclosure-quality–volatility relationship.

Third, an enhanced SASB metric should be developed that captures quantitative disclosure (greenhouse-gas figures, financed-emissions calculations, scenario-analysis numbers) in addition to textual coverage. Structured-data parsing of report tables, supplemented by manual classification of disclosed metrics, would distinguish banks that genuinely engage with material topics from those that merely mention them. The null result on H3 in the present analysis may reflect the inability of keyword-density metrics to capture this distinction.

Fourth, the selective-obfuscation finding documented in Sections 5.1 and 5.5 invites several refinements. A paragraph-level rather than sentence-level material-text identification would capture material-topic discussion that extends across multiple sentences without uniform keyword repetition; analysing the gap by SASB topic separately would identify whether obfuscation is concentrated in particular material topics (for example, financed emissions); and an event-study design built around regulatory enforcement actions or controversial

disclosures would test whether selective obfuscation increases in periods of pending material adverse news.

Fifth, language-adjusted analysis would extend the sample to non-English ESG reports, particularly for Chinese, French and German banks where substantial sustainability disclosure occurs in the local language. A multi-language extension of the L-M sentiment dictionary, combined with translated SASB keyword sets, would enable a true global panel analysis rather than the English-language-dominated sample examined here.

Sixth, individual-investor and analyst-attention data would help disentangle the mechanisms through which disclosure quality affects market outcomes. If the readability channel operates by slowing processing of bad news, then the obfuscation effect should be sharper in retail-investor-heavy stocks than in institutional-investor-heavy ones, and should be moderated by analyst coverage.

6.7 Conclusion

This thesis has examined whether the quality of bank ESG disclosures, measured through four text-analytic dimensions on a global panel of 80 banks across the 2020–2024 period, predicts stock return volatility. The headline empirical result is that two stylistic dimensions of disclosure — linguistic credibility, measured by tone, and linguistic accessibility, measured by readability — are robustly associated with volatility, while the substantive content dimensions measured by ESG and SASB density are not. The findings support a partial version of the information asymmetry framework, in which the manner of disclosure shapes investor uncertainty independently of the substantive performance the disclosure ostensibly describes.

The companion test of the impression management hypothesis reveals a now-sharper interpretive asymmetry. Managers respond to weak financial performance by obscuring information through linguistic complexity (the obfuscation channel, supported at $p = 0.028$), but the data do not support the tone-inflation channel in either direction. This pattern suggests that the two stylistic dimensions of disclosure carry distinct strategic logics: readability tracks contemporaneous performance, tone does not.

A separate descriptive contribution is the documentation of a near-universal pattern of selective obfuscation of material content. In 99.5% of the reports in the sample, the Fog Index of sentences referencing SASB material topics is higher than the Fog Index of the rest of the report; in 91.3% of reports, the L-M Tone of material content is less positive than the tone of non-material content. Both effects are very large by conventional standards. This is a structural feature of bank ESG reporting that has not previously been quantified, with direct implications for the design of sustainability disclosure standards.

For the design of sustainability disclosure frameworks under the CSRD, ESRS and ISSB, the findings argue for pairing content requirements with explicit quality standards — particularly readability standards, and ideally readability standards applied specifically to material-topic discussion — to ensure that mandated coverage translates into accessible, actionable information for investors. The evolution from voluntary to mandatory sustainability disclosure that is now underway across jurisdictions creates both an opportunity and a research imperative: an opportunity for regulators to design the next generation of disclosure rules with explicit attention to communicative quality, and an imperative for researchers to track the impact of these rules using carefully constructed, multi-dimensional measures of disclosure quality of the kind developed in this thesis.

References

- Acharya, A., Blackwell, M., & Sen, M. (2016). Explaining causal findings without bias: Detecting and assessing direct effects. *American Political Science Review*, 110(3), 512–529.
- Akerlof, G. A. (1970). The market for “lemons”: Quality uncertainty and the market mechanism. *Quarterly Journal of Economics*, 84(3), 488–500.
- Baier, P., Berninger, M., & Kiesel, F. (2020). Environmental, social and governance reporting in annual reports: A textual analysis. *Financial Markets, Institutions & Instruments*, 29(3), 93–118.
- Beck, N., & Katz, J. N. (1995). What to do (and not to do) with time-series cross-section data. *American Political Science Review*, 89(3), 634–647.
- Bloomfield, R. (2002). The 'incomplete revelation hypothesis' and financial reporting. *Accounting Horizons*, 16(3), 233–243.
- Brennan, N. M., Guillamon-Saorin, E., & Pierce, A. (2009). Impression management: Developing and illustrating a scheme of analysis for narrative disclosures. *Accounting, Auditing & Accountability Journal*, 22(5), 789–832.
- Caglio, A., Melloni, G., & Perego, P. (2020). Informational content and assurance of textual disclosures: Evidence on integrated reporting. *European Accounting Review*, 29(1), 55–83.
- Clarkson, P. M., Li, Y., Richardson, G. D., & Vasvari, F. P. (2008). Revisiting the relation between environmental performance and environmental disclosure: An empirical analysis. *Accounting, Organizations and Society*, 33(4–5), 303–327.
- Clatworthy, M., & Jones, M. J. (2003). Financial reporting of good news and bad news: Evidence from accounting narratives. *Accounting and Business Research*, 33(3), 171–185.
- Courtis, J. K. (1998). Annual report readability variability: Tests of the obfuscation hypothesis. *Accounting, Auditing & Accountability Journal*, 11(4), 459–472.
- Diamond, D. W., & Verrecchia, R. E. (1991). Disclosure, liquidity, and the cost of capital. *Journal of Finance*, 46(4), 1325–1359.
- Drempetic, S., Klein, C., & Zwergel, B. (2020). The influence of firm size on the ESG score: Corporate sustainability ratings under review. *Journal of Business Ethics*, 167(2), 333–360.

- Duan, Y., Li, R., & Xue, J. (2022). Bank stock volatility during the COVID-19 pandemic. *Finance Research Letters*, 47, 102772.
- El-Haj, M., Rayson, P., Walker, M., Young, S., & Simaki, V. (2019). In search of meaning: Lessons, resources and next steps for computational analysis of financial discourse. *Journal of Business Finance & Accounting*, 46(3–4), 265–306.
- Flannery, M. J., & Thakor, A. V. (2006). Accounting, transparency and bank stability. *Journal of Financial Intermediation*, 15(3), 281–284.
- Friede, G., Busse, M., & Bassen, A. (2015). ESG and financial performance: Aggregated evidence from more than 2000 empirical studies. *Journal of Sustainable Finance & Investment*, 5(4), 210–233.
- Grossman, S. J. (1981). The informational role of warranties and private disclosure about product quality. *Journal of Law and Economics*, 24(3), 461–483.
- Gunning, R. (1952). *The technique of clear writing*. New York, NY: McGraw-Hill.
- Healy, P. M., & Palepu, K. G. (2001). Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. *Journal of Accounting and Economics*, 31(1–3), 405–440.
- Huang, X., Teoh, S. H., & Zhang, Y. (2014). Tone management. *The Accounting Review*, 89(3), 1083–1113.
- Khan, M., Serafeim, G., & Yoon, A. (2016). Corporate sustainability: First evidence on materiality. *The Accounting Review*, 91(6), 1697–1724.
- Lehavy, R., Li, F., & Merkley, K. (2011). The effect of annual report readability on analyst following and the properties of their earnings forecasts. *The Accounting Review*, 86(3), 1087–1115.
- Leuz, C., & Verrecchia, R. E. (2000). The economic consequences of increased disclosure. *Journal of Accounting Research*, 38, 91–124.
- Li, F. (2008). Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics*, 45(2–3), 221–247.

- 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.
- Lo, K., Ramos, F., & Rogo, R. (2017). Earnings management and annual report readability. *Journal of Accounting and Economics*, 63(1), 1–25.
- Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *Journal of Finance*, 66(1), 35–65.
- Loughran, T., & McDonald, B. (2016). Textual analysis in accounting and finance: A survey. *Journal of Accounting Research*, 54(4), 1187–1230.
- Lyon, T. P., & Maxwell, J. W. (2011). Greenwash: Corporate environmental disclosure under threat of audit. *Journal of Economics & Management Strategy*, 20(1), 3–41.
- Marquis, C., & Toffel, M. W. (2014). Scrutiny, norms, and selective disclosure: A global study of greenwashing. *Organization Science*, 27(2), 483–504.
- Melloni, G., Caglio, A., & Perego, P. (2017). Saying more with less? Disclosure conciseness, completeness and balance in integrated reports. *Journal of Accounting and Public Policy*, 36(3), 220–238.
- Milgrom, P. R. (1981). Good news and bad news: Representation theorems and applications. *The Bell Journal of Economics*, 12(2), 380–391.
- Morgan, D. P. (2002). Rating banks: Risk and uncertainty in an opaque industry. *American Economic Review*, 92(4), 874–888.
- Nier, E., & Baumann, U. (2006). Market discipline, disclosure and moral hazard in banking. *Journal of Financial Intermediation*, 15(3), 332–361.
- Price, S. M., Doran, J. S., Peterson, D. R., & Bliss, B. A. (2012). Earnings conference calls and stock returns: The incremental informativeness of textual tone. *Journal of Banking & Finance*, 36(4), 992–1011.
- Sustainability Accounting Standards Board. (2023). Commercial Banks Sustainability Accounting Standard (FN-CB), Version 2023-12.
- Spence, M. (1973). Job market signaling. *Quarterly Journal of Economics*, 87(3), 355–374.

Zhou, X. (2001). Understanding the determinants of managerial ownership and the link between ownership and performance: Comment. *Journal of Financial Economics*, 62(3), 559–571.

Appendix A: “Sample Banks”

The final analytical sample comprises 80 listed commercial banks across five global regions, yielding 390 bank-year observations over the period 2020–2024. Banks are drawn from the S&P Global Top 100 Banks by Total Assets ranking (December 2024 edition). Each bank listed below contributes the full five-year panel (2020–2024) unless a restricted year range is indicated; restricted ranges reflect the absence of an English-language standalone or integrated sustainability report for the year(s) not listed.

Note: The Capital One 2020 ESG report is not included in the panel because the file is a combined 2019–2020 disclosure whose content straddles a year outside the 2020–2024 sampling window. Capital One is therefore present in the panel for 2021–2023 only.

Asia-Pacific (38 banks)

Bank	Country	Ticker	Report years
ANZ Group Holdings	Australia	ANZ.AX	2020–2024
Agricultural Bank of China	China	1288.HK	2020–2024
Bank of China	China	3988.HK	2020–2024
Bank of Communications	China	3328.HK	2020–2024
Bank of Jiangsu	China	600919.SS	2021–2024
China Construction Bank	China	0939.HK	2020–2024
China Everbright Bank	China	601818.SS	2020–2024
China Merchants Bank	China	3968.HK	2020–2024
China Minsheng Banking Corp	China	1988.HK	2020–2024
China Zheshang Bank	China	2016.HK	2020–2024
CITIC Limited	Hong Kong	0267.HK	2020–2024
Commonwealth Bank of Australia	Australia	CBA.AX	2020–2024
DBS Group Holdings	Singapore	D05.SI	2020–2024
Hana Financial Group	South Korea	086790.KS	2020–2024
HDFC Bank	India	HDFCBANK.NS	2020–2023
HSBC Holdings	UK	HSBA.L	2020–2024
Huaxia Bank	China	600015.SS	2020–2024
Industrial and Commercial Bank of China	China	1398.HK	2020–2024
Industrial Bank (CIB)	China	601166.SS	2020–2024
Japan Post Bank	Japan	7182.T	2020–2022, 2024
KB Financial Group	South Korea	105560.KS	2020–2024
Mitsubishi UFJ Financial Group (MUFG)	Japan	8306.T	2020–2024
Mizuho Financial Group	Japan	8411.T	2020–2024
National Australia Bank	Australia	NAB.AX	2020–2024
Nomura Holdings	Japan	8604.T	2020–2024

OCBC Group	Singapore	O39.SI	2020–2024
Ping An Insurance Group	China	2318.HK	2020–2024
Postal Savings Bank of China	China	1658.HK	2020–2024
Resona Holdings	Japan	8308.T	2020–2024
Shanghai Pudong Development Bank	China	600000.SS	2020–2024
Shinhan Financial Group	South Korea	055550.KS	2020–2024
Standard Chartered	UK	STAN.L	2020–2024
State Bank of India	India	SBIN.NS	2020–2024
Sumitomo Mitsui Banking Corp (SMBC)	Japan	8316.T	2020–2024
Sumitomo Mitsui Trust Holdings	Japan	8309.T	2020–2024
United Overseas Bank	Singapore	U11.SI	2020–2024
Westpac Banking Corporation	Australia	WBC.AX	2020–2024
Woori Financial Group	South Korea	316140.KS	2020–2024

Europe (20 banks)

Bank	Country	Ticker	Report years
ABN AMRO Bank	Netherlands	ABN.AS	2020–2024
Banco Bilbao Vizcaya Argentaria	Spain	BBVA.MC	2020–2024
Banco Santander	Spain	SAN.MC	2020–2024
Barclays	UK	BARC.L	2020, 2022, 2024
BNP Paribas	France	BNP.PA	2020–2024
CaixaBank	Spain	CABK.MC	2020–2024
Commerzbank	Germany	CBK.DE	2020–2024
Crédit Agricole Group	France	ACA.PA	2020–2024
Danske Bank	Denmark	DANSKE.CO	2020–2024
Deutsche Bank	Germany	DB	2020–2024
Erste Group Bank	Austria	EBS.VI	2020–2024
Intesa Sanpaolo	Italy	ISP.MI	2020–2024
KBC Group	Belgium	KBC.BR	2020–2024
Lloyds Banking Group	UK	LLOY.L	2020–2024
NatWest Group	UK	NWG.L	2020–2024
Nordea Bank	Finland	NDA-FI.HE	2020–2024
SEB Group	Sweden	SEB-A.ST	2020–2024
Société Générale	France	GLE.PA	2020–2024
UBS Group	Switzerland	UBSG.SW	2020–2024
UniCredit	Italy	UCG.MI	2020–2024

North America (17 banks)

Bank	Country	Ticker	Report years
Bank of New York Mellon	USA	BK	2020–2024
Capital One Financial Corp.	USA	COF	2021–2023

Charles Schwab Corp.	USA	SCHW	2021–2023
CIBC (Canadian Imperial Bank of Commerce)	Canada	CM.TO	2020–2024
Citigroup	USA	C	2020–2024
Goldman Sachs	USA	GS	2020–2024
JPMorgan Chase	USA	JPM	2020–2024
Morgan Stanley	USA	MS	2020–2024
National Bank of Canada	Canada	NA.TO	2020–2024
PNC Financial Services Group	USA	PNC	2021–2024
Royal Bank of Canada	Canada	RY.TO	2020–2024
Scotiabank (Bank of Nova Scotia)	Canada	BNS.TO	2020–2024
State Street	USA	STT	2020–2024
TD Bank Group	Canada	TD.TO	2020–2024
Truist Financial	USA	TFC	2020–2024
U.S. Bancorp	USA	USB	2020–2024
Wells Fargo	USA	WFC	2020–2024

Latin America (3 banks)

Bank	Country	Ticker	Report years
Banco Bradesco	Brazil	BBDC4.SA	2020–2024
Banco do Brasil	Brazil	BBAS3.SA	2020–2024
Itaú Unibanco Holding	Brazil	ITUB4.SA	2020–2024

Middle East (2 banks)

Bank	Country	Ticker	Report years
First Abu Dhabi Bank	UAE	FAB	2020–2024
Qatar National Bank (QNB Group)	Qatar	QNBK.QA	2020–2024

Unlisted institutions analysed but absent from the regression panel

The following cooperative, mutual or state-owned institutions are processed through the text-analytic pipeline but cannot enter the regression panel because the dependent variable (annualised stock return volatility) cannot be computed: BPCE (Groupe BPCE), La Banque Postale, Raiffeisen Switzerland, Crédit Mutuel, Norinchukin Bank, Rabobank, DZ Bank, and Landesbank Baden-Württemberg (LBBW). Their text-analytic metrics are retained in the underlying dataset for completeness.

Appendix B: “SASB Material Topic Dictionary”

The SASB material-topic dictionary used in this thesis covers the six material topics identified by the Sustainability Accounting Standards Board for the Commercial Banking industry (SASB Standard FN-CB, Version 2023-12). Each topic is operationalised through a list of primary keywords (high-signal terms drawn directly from the SASB standard) and secondary keywords (contextual terms common in banking ESG disclosure). The dictionary contains 124 primary and 203 secondary keywords, for a total of 327 terms.

Data Security (FN-CB-230a)

SASB Code: FN-CB-230a

Primary keywords (21): data breach, data breaches, data security, data privacy, data protection, cybersecurity, cyber security, cyber risk, cyber threat, cyber attack, information security, personal data, account holder, data loss, identity theft, phishing, ransomware, malware, encryption, gdpr, data governance

Secondary keywords (28): privacy, firewall, intrusion, vulnerability, authentication, biometric, tokenization, penetration testing, security incident, security awareness, security training, cloud security, digital security, fraud prevention, fraud detection, customer data, confidentiality, unauthorized access, two-factor, multi-factor, access control, security operations, threat intelligence, security breach, incident response, data handling, information protection, pci dss

Financial Inclusion and Capacity Building (FN-CB-240a)

SASB Code: FN-CB-240a

Primary keywords (16): financial inclusion, financial literacy, financial education, unbanked, underbanked, underserved, community development, small business lending, community lending, affordable banking, capacity building, community reinvestment, microfinance, microlending, financial access, banking access

Secondary keywords (29): low-income, rural banking, mobile banking, digital banking, financial empowerment, economic empowerment, community investment, social impact,

affordable housing, community development financial, minority-owned, women-owned, small enterprise, financial capability, savings program, checking account, community outreach, economic development, inclusive finance, developing countries, developing economies, emerging markets, poverty alleviation, poverty reduction, sme lending, sme financing, disadvantaged, vulnerable communities, accessibility

Incorporation of ESG Factors in Credit Analysis (FN-CB-410a)

SASB Code: FN-CB-410a

Primary keywords (20): esg integration, esg factors in credit, esg credit analysis, esg risk assessment, esg due diligence, esg screening, sustainable lending, sustainable finance, green lending, green bond, green bonds, social bond, social bonds, sustainability-linked loan, sustainability-linked bond, responsible lending, responsible investment, climate risk assessment, environmental risk assessment, transition finance

Secondary keywords (34): esg criteria, esg policy, esg framework, esg scoring, environmental risk, social risk, governance risk, credit risk management, sustainability risk, equator principles, principles for responsible, green finance, climate finance, social finance, impact investing, impact investment, responsible banking, sustainable development goals, sdg, sdgs, taxonomy, eu taxonomy, green taxonomy, exclusion policy, sector policy, sensitive sector, controversial weapons, coal financing, fossil fuel financing, positive screening, negative screening, best-in-class, thematic investing, blended finance

Financed Emissions and Climate Risk (FN-CB-410b)

SASB Code: FN-CB-410b

Primary keywords (22): financed emissions, scope 1, scope 2, scope 3, carbon footprint, carbon emissions, greenhouse gas, ghg emissions, co2 emissions, net zero, net-zero, carbon neutral, carbon neutrality, climate target, science based targets, sbti, paris agreement, paris-aligned, tcf, climate-related financial disclosures, decarbonization, decarbonisation

Secondary keywords (37): carbon intensity, emission reduction, emissions reduction, climate change, climate risk, climate strategy, transition risk, physical risk, stranded assets, carbon pricing, carbon offset, carbon credit, renewable energy, energy efficiency, clean

energy, climate scenario, scenario analysis, stress testing climate, 2 degrees, 1.5 degrees, global warming, fossil fuel, coal phase, oil and gas, pcaf, partnership for carbon accounting, carbon disclosure, cdp, climate disclosure, green building, energy transition, low carbon, low-carbon, climate action, climate commitment, environmental footprint, operational emissions

Business Ethics (FN-CB-510a)

SASB Code: FN-CB-510a

Primary keywords (25): business ethics, whistleblower, whistleblowing, anti-corruption, anticorruption, bribery, anti-bribery, insider trading, market manipulation, antitrust, anti-competitive, anticompetitive, money laundering, anti-money laundering, aml, know your customer, kyc, compliance, code of conduct, code of ethics, fraud, misconduct, malpractice, sanctions, financial crime

Secondary keywords (35): regulatory compliance, regulatory fine, legal proceedings, ethical conduct, ethical behavior, ethical behaviour, integrity, transparency, accountability, conflict of interest, conflicts of interest, fair dealing, fair treatment, treating customers fairly, tax transparency, tax compliance, tax evasion, corruption, speak up, reporting mechanism, ethics training, compliance training, regulatory enforcement, regulatory action, responsible marketing, responsible sales, customer protection, consumer protection, mis-selling, mis selling, product governance, politically exposed, suspicious transaction, beneficial ownership, terrorist financing

Systemic Risk Management (FN-CB-550a)

SASB Code: FN-CB-550a

Primary keywords (20): systemic risk, systemically important, g-sib, gsib, too big to fail, stress test, stress testing, capital adequacy, capital requirements, capital planning, basel iii, basel iv, basel framework, resolution planning, living will, recovery plan, macroprudential, financial stability, total loss-absorbing capacity, tlac

Secondary keywords (40): capital ratio, tier 1, tier 2, cet1, leverage ratio, liquidity ratio, lcr, nsfr, risk management, risk governance, risk appetite, risk framework, enterprise risk,

operational risk, credit risk, market risk, liquidity risk, counterparty risk, concentration risk, regulatory capital, risk-weighted assets, rwa, contingency planning, business continuity, dodd-frank, volcker rule, mifid, prudential regulation, prudential supervision, central bank, regulator, supervisory review, interconnectedness, complexity, cross-jurisdictional, pillar 3, disclosure requirements, buffer, countercyclical, capital buffer

Appendix C: “Variable Definitions”

This appendix defines every variable used in the regression analyses reported in Chapter 5, organised by role. All continuous variables are winsorised at the 1st and 99th percentiles in estimation. Year-varying financial controls are computed at the calendar-year frequency from daily price data (volatility, beta) or year-end balance-sheet values (market capitalisation, total assets, ROE).

Dependent variable

Variable	Definition
<i>annualized_volatility</i>	Stock-return volatility for bank <i>i</i> in calendar year <i>t</i> . Computed as the standard deviation of daily log returns over year <i>t</i> , annualised by multiplying by $\sqrt{252}$. A minimum of 100 daily observations per bank-year is required.

Disclosure-quality independent variables

Variable	Definition
<i>fog_index</i>	Gunning Fog Index of the bank's ESG report for year <i>t</i> . Computed as $0.4 \times (\text{avg. words per sentence} + \% \text{ complex words})$, where complex words have ≥ 3 syllables. Higher values indicate less readable text.
<i>flesch_kincaid_grade</i>	Flesch-Kincaid Grade Level of the report, used as a Fog-Index robustness check in R2. Computed as $0.39 \times (\text{avg. words per sentence}) + 11.8 \times (\text{avg. syllables per word}) - 15.59$.
<i>lm_tone</i>	Loughran-McDonald net positive sentiment ratio, $(P - N) / (P + N)$, where <i>P</i> and <i>N</i> are positive- and negative-word counts using the L-M 2011 financial-sentiment dictionary. Range $[-1, +1]$; higher values indicate more positive tone.
<i>abs_tone</i>	Absolute value of <i>lm_tone</i> , used in §5.3 as a measure of tone extremity independent of direction.
<i>esg_density_pct</i>	ESG vocabulary density: count of ESG-related words from the Baier et al. (2020) banking-adapted lexicon divided by total clean word count, expressed as a percentage.
<i>sasb_density_pct</i>	SASB material-topic density: total weighted hits across the six SASB Commercial Banking material topics divided by clean word count, expressed as a percentage. Primary keywords are weighted twice; secondary keywords once.
<i>sasb_hhi_balance</i>	Normalised Herfindahl-Hirschman balance score across the six SASB material topics: $1 - (\text{HHI} - 1/N) / (1 - 1/N)$, where HHI is computed on each topic's share of total SASB hits and <i>N</i> = 6. Range $[0, 1]$; 1 indicates perfectly even coverage across all six topics, 0 indicates concentration on a single topic.
<i>sasb_proportional_coverage</i>	Proportional-coverage metric: percentage of the six SASB topics for which weighted hits exceed 0.1% of the bank's clean word count. Used as an alternative SASB measure in Model 1e.

Topic-conditioned disclosure-quality variables (Tables 8–9)

Variable	Definition
<i>fog_material</i>	Gunning Fog Index computed only on sentences containing at least one SASB primary-keyword match. Captures the readability of the report's material-topic discussion.
<i>fog_non_material</i>	Gunning Fog Index computed on the complement: sentences with no SASB primary-keyword match.
<i>fog_gap</i>	<i>fog_material</i> – <i>fog_non_material</i> . Within-document gap measure; positive values indicate that material content is less readable than the rest of the report (selective obfuscation).

<i>tone_material</i>	Loughran-McDonald tone computed only on sentences containing at least one SASB primary-keyword match.
<i>tone_non_material</i>	Loughran-McDonald tone computed on the complement set of sentences.
<i>tone_gap</i>	<i>tone_material</i> – <i>tone_non_material</i> . Within-document gap; negative values indicate that material content is less positively toned than the rest of the report.

Financial and report-level controls

Variable	Definition
<i>log_market_cap</i>	Natural logarithm of year-end market capitalisation in US dollars. Bank-year market capitalisations are sourced from primary annual reports and third-party financial databases, with year-end currency conversion to USD.
<i>roe</i>	Return on equity for year <i>t</i> : net income divided by year-end total stockholders' equity. Sourced bank-by-bank from primary annual reports.
<i>beta</i>	Year-varying market beta: slope coefficient from a regression of the bank's daily log returns on daily log returns of the iShares MSCI ACWI ETF for the corresponding calendar year. A minimum of 100 overlapping trading days per year is required.
<i>total_assets</i>	Year-end total assets in USD billions, sourced from each bank's audited primary financial statements with year-end currency conversion (see methodology notes).
<i>log_total_assets</i>	Natural logarithm of <i>total_assets</i> . Used as an explicit bank-size control in robustness check R5.
<i>report_dummy</i>	Binary indicator: 0 = standalone ESG/CSR/Sustainability report; 1 = integrated annual report with embedded sustainability content.
<i>clean_word_count</i>	Number of cleaned words in the report after extraction-pipeline preprocessing (removal of repeated sidebar text, headers, footers, URLs and non-ASCII artefacts).
<i>log_word_count</i>	Natural logarithm of <i>clean_word_count</i> . Included as a control to absorb report-length effects.
<i>tier1_ratio</i>	The ratio of the bank's Tier 1 (going-concern) regulatory capital to risk-weighted assets, the supervisory measure of loss-absorbing capital strength under Basel III. The fully-loaded ratio is used where disclosed, the transitional ratio where the fully-loaded figure is unavailable (held consistent per bank across years), and the standardised-approach ratio for US banks. Banks with a 31 March fiscal year-end (the Japanese banks and HDFC Bank) are dated to that fiscal year-end. Used as a solvency control in the augmented specification (Model 1b+T1) and decomposed in the Mundlak analysis.

Fixed effects

Variable	Definition
<i>year_2021, year_2022, year_2023, year_2024</i>	Year fixed-effect dummies; 2020 is the omitted base category. Absorb common time shocks including the COVID-19 baseline.
<i>region_Europe, region_Latin_America, region_Middle_East, region_North_America</i>	Region fixed-effect dummies; Asia-Pacific is the omitted base category. Absorb systematic regional differences in disclosure norms, regulatory environment, and bank-business-model composition.

Estimators (Table 7)

Variable	Definition
<i>Pooled OLS with bank-clustered SE</i>	Baseline specification used throughout Chapter 5. Standard errors are clustered at the bank (<i>folder_name</i>) level to accommodate arbitrary within-bank serial correlation in residuals while assuming independence across banks.

<i>Beck-Katz Panel-Corrected Standard Errors (PCSE)</i>	Beck & Katz (1995). Pooled OLS coefficients with a sandwich-form variance estimator that allows panel-specific heteroskedasticity and contemporaneous cross-panel correlation in residuals. Treated as the canonical robustness check for time-series cross-section panels in which year fixed effects do not exhaust common time-shock structure.
<i>Feasible GLS with AR(1) errors</i>	Prais-Winsten transformation under the assumption $\varepsilon_{it} = \rho \varepsilon_{i,t-1} + u_{it}$ with a common ρ estimated from pooled within-panel residuals (estimated $\rho = 0.45$ in this sample). Yields new and more efficient coefficient estimates if the AR(1) structure is correctly specified.
<i>Feasible GLS with panel heteroskedasticity</i>	Weighted least squares with weights equal to $1/\hat{\sigma}_i^2$, where $\hat{\sigma}_i^2$ is the sample variance of bank i 's residuals from a first-stage OLS. Permits each bank to have its own residual variance but imposes no serial-correlation structure.
<i>Bank Fixed Effects</i>	Within-bank estimator. Region dummies are absorbed by the bank fixed effects. The F-test for poolability rejects pooling at the 1% level ($F = 4.61$), confirming that bank-level heterogeneity is substantively important; however, the short five-year panel yields limited within-bank time-series variation in the persistent disclosure-quality variables of interest.
<i>Random Effects</i>	Random-effects GLS estimator (linearmodels.RandomEffects). Provides an intermediate specification between pooled OLS and bank fixed effects, retaining a portion of cross-bank identifying variation while accounting for bank-level random heterogeneity. Used in §5.5 alongside bank fixed effects as a Hausman comparison.

Appendix D: “Methodology Notes Register”

This appendix documents the year-end USD reference rates used to convert balance-sheet and market-capitalisation values denominated in non-US currencies. Rates are sourced via the yfinance API (FX pair tickers EURUSD=X, USDKRW=X, USDCAD=X, USDJPY=X, etc.) at the close of 31 December of each calendar year for all banks except those whose fiscal year ends 31 March, for which the corresponding 31 March rate is used. The four most-frequent conversions are tabulated below; conversions for other currencies (GBP, CHF, HKD, INR, AED, QAR, AUD, DKK, SEK, SGD) follow the same convention and are available in the underlying dataset.

Currency pair	2020	2021	2022	2023	2024
EUR/USD (31 Dec)	1.2300	1.1325	1.0661	1.1068	1.0406
KRW/USD (31 Dec)	0.000920	0.000840	0.000793	0.000783	0.000681
CAD/USD (31 Dec)	0.7843	0.7847	0.7379	0.7559	0.6969
JPY/USD (31 Mar)	0.009256	0.009066	0.008195	0.007507	0.006603

JPY/USD rates are reported at 31 March because Mizuho Financial Group, the only Japanese bank in the panel whose published total-asset figures correspond to the 31 March fiscal year-end, accounts for all Japanese-yen conversions in the bank-size variable.

Author's Statement:

I hereby expressly declare that, according to the article 8 of Law 1559/1986, this dissertation is solely the product of my personal work, does not infringe any intellectual property, personality and personal data rights of third parties, does not contain works/contributions from third parties for which the permission of the authors/beneficiaries is required, is not the product of partial or total plagiarism, and that the sources used are limited to the literature references alone and meet the rules of scientific citations.