Essays on Information Processing in Empirical Finance

Inauguraldissertation
zur
Erlangung des Doktorgrades
der Wirtschafts- und Sozialwissenschaftlichen Fakultät
der
Universität zu Köln

vorgelegt
von

M.Sc. Nathalie Zährl
aus
Marburg

Referent: Prof. Dr. Dieter Hess

Korreferent: Jun. Prof. Dr. Tobias Bauckloh

Tag der Promotion: 15.07.2026

Acknowledgements

This dissertation was written during my time at the University of Cologne, and its completion would not have been possible without the support of many people who accompanied me throughout this journey.

First and foremost, I would like to express my sincere gratitude to my supervisor, Prof. Dr. Dieter Hess, for his continuous guidance, encouragement, and thoughtful feedback throughout my doctoral studies. His high standards and constructive criticism significantly improved this dissertation, and I am very grateful for his support far beyond the written work. I am also deeply grateful to my co-supervisor, Jun.-Prof. Dr. Tobias Bauckloh, for his valuable advice, encouragement, and continuous support throughout this process.

I owe special thanks to my co-author, Luca Conrads, for the many conversations and the steady support along the way. You made this project, and my time at the chair, much brighter. I am also grateful to my colleagues at the Department of Finance for their feedback, support, and the many conversations that made this time both productive and enjoyable.

Beyond academia, I am grateful to my friends for their companionship and encouragement throughout the ups and downs of this process. Thank you for listening, for making me laugh when I needed it most, and for reminding me that there is always more than work.

Finally, I want to thank my family with all my heart. Your unwavering love and support have been the constant foundation beneath everything I have done. I truly could not have done this without you.

Most of all, I thank Christian Hinz. Thank you for your patience, your steady encouragement, and for being there through the demanding phases, the doubts, and the long days. Your love and support meant more to me than I can put into words, and I will always be grateful for having you by my side.

Nathalie Zährl

Contents

Acknowledgements **I**

List of Tables **VII**

List of Figures **VIII**

1 Introduction **1**

2 The EPS Growth Illusion: How Analysts' Share Count Forecasts Shape Valuation Perceptions **7**

2.1 Introduction 7

2.2 Theoretical background 11

2.3 Data 14

2.4 Results 15

2.4.1 Discrepancies in earnings forecasts 16

2.4.2 Impact on forward price–earnings ratio 20

2.4.3 Impact on implied cost of capital 22

2.4.4 Share repurchases and issuances: firm characteristics and forecast accuracy 31

2.5 Conclusion 39

3 From Street to GAAP: Forecast Definitions in Implied Cost of Capital Estimation **42**

3.1 Introduction 42

3.2 Theoretical background 46

3.3 Data and methodology 49

3.3.1 Data and sample construction 49

3.3.2 Methodology 50

3.4 Discrepancies in earnings forecasts and ICCs . . . 52
3.4.1 Forecast differences: levels and implied growth . . . 53
3.4.2 From forecast differences to ICC differences . . . 55
3.4.3 Implications for ICC rankings . . . 60
3.5 Reliability and performance of ICCs . . . 61
3.5.1 Measurement-error variance of Street- and GAAP-based ICCs . . . 61
3.5.2 Street versus GAAP ICC . . . 63
3.5.3 GAAP ICC and the Street–GAAP ICC differential . . . 64
3.5.4 High earnings-divergence states . . . 65
3.5.5 Cross-sectional heterogeneity . . . 66
3.5.6 Model-level evidence . . . 68
3.6 Conclusion . . . 72

4 Beyond ESG Scores: Risk and Return Insights from SFDR Metrics 74
4.1 Introduction . . . 74
4.2 Related literature . . . 76
4.3 Data and methodology . . . 79
4.3.1 Data . . . 79
4.3.2 Score calculation . . . 79
4.3.3 Descriptive statistics . . . 80
4.4 Financial performance . . . 82
4.4.1 Returns across sustainability deciles . . . 83
4.4.2 Investment strategies . . . 87
4.5 Conclusion . . . 97

A Appendix to Chapter 2 99
A.1 Variable definitions . . . 99
A.2 Implied cost of capital models . . . 103

B Appendix to Chapter 3 105
B.1 Implied cost of capital models . . . 105
B.2 Measurement-error-variance diagnostics for expected-return proxies . . . 107

C Appendix to Chapter 4 110
C.1 Variable definitions . . . 110

Bibliography **112**

Lebenslauf **120**

Eidesstattliche Erklärung **121**

List of Tables

2.1 Availability of LSEG forecast variables 15
2.2 Forecasted number of shares growth 17
2.3 Earnings forecasts and differences by share growth deciles 19
2.4 Median growth rates implied in earnings forecasts 20
2.5 Forward P/E ratios: Summary statistics and median differences 22
2.6 Summary statistics of ICC estimates 24
2.7 Median differences between EPS- and FL-based ICC estimates by decile . 26
2.8 Mean differences in ICCs by decile 29
2.9 Measurement error diagnostics 30
2.10 Historical and forecasted growth in number of shares outstanding 32
2.11 Firm characteristics across deciles sorted by changes in shares outstanding 33
2.12 Spearman rank correlation matrices: actual and forecasted data 34
2.13 Regression results of forecasted and actual change in shares outstanding 36
2.14 Forecast vs. actual share changes: regression of shares and forecast errors 38
2.15 Summary statistics of forecast errors by type and horizon 39
2.16 Mean differences in forecast errors 40

3.1 Sample and analyst forecast availability 50
3.2 Summary statistics of Street and GAAP earnings forecasts 53
3.3 Implied growth in analyst's forecast under Street and GAAP definitions . 54
3.4 Sequential forecast growth rates: EPS vs. GAAP 54
3.5 Implied cost of capital levels by valuation model: Street versus GAAP forecast inputs 56
3.6 Sensitivity of ICC models to earnings choice: Street–GAAP ICC difference 59
3.7 Rank correlations and tail turnover of Street- and GAAP-based ICCs . . . 60

3.8 Measurement-error variance by ICC model: Street versus GAAP 62
3.9 Baseline return regressions: Street versus GAAP composite ICC 64
3.10 Return regressions with GAAP ICC and the Street–GAAP ICC differential across horizons 65
3.11 High earnings-divergence states: implied marginal effects 66
3.12 GAAP ICC and the Street–GAAP ICC differential across firm size terciles 67
3.13 Industry heterogeneity: GAAP ICC and the Street–GAAP ICC differential 68
3.14 Return regressions by ICC model: Street versus GAAP 69
3.15 Return regressions by ICC model: GAAP ICC and the Street–GAAP ICC differential across horizons 70
3.16 Model-level size heterogeneity 71
3.17 Model-level industry heterogeneity 71

4.1 Descriptive statistics 81
4.2 Correlation matrix 82
4.3 Mean score per decile 84
4.4 Mean return per decile 86
4.5 Annual mean returns 89
4.6 Annual mean return spreads 91
4.7 Annual mean return differences 92
4.8 Sharpe ratios 93
4.9 Three-year annualized mean returns and sharpe ratios 94
4.10 Three-year annualized mean return spreads 95

A.1 Definitions and formulas of key variables 99
A.2 Implied cost of capital models 104

B.1 Implied cost of capital models 106

C.1 PAI variables used in the SFDR Score construction 111

List of Figures

2.1 Forecasted growth in outstanding shares 18
2.2 Average ICC estimates across models 25
2.3 Differences in composite ICC 26
2.4 Differences in ICC estimates across forecast types per model 27

3.1 ICC differences by level and growth divergence deciles 58
3.2 Distribution of composite ICC differences from 2004 to 2024 59

4.1 Overlap of firms across deciles 88

Chapter 1

Introduction

In capital markets research, the objects we can measure are rarely the objects we ultimately care about. Expected profitability and growth, required returns, and sustainability-related risk are typically observed only through summary signals that compress information into a single number. These summaries are attractive because they make complex information comparable across firms and usable in valuation, trading, and empirical tests. They also shape the language of markets: growth narratives are often told in per-share terms, valuation comparisons are expressed through multiples, expected returns are proxied by implied discount rates, and sustainability is increasingly summarized by a score. Viewed this way, many empirical inputs are outcomes of *information processing*: rules that translate rich underlying information into simple, comparable signals.

However, the same convenience creates a basic tension. A summary signal is only as informative as the rules used to construct it. Even small, seemingly routine construction choices can introduce mechanical components that alter the signal in economically meaningful ways. These components are features of the measurement architecture rather than underlying fundamentals. This is important because much of the evidence in accounting and finance is based on rankings, tails, and portfolio sorts. Once a measure is used to classify firms, even minor mechanical shifts can alter which firms fall into the "cheap," "high expected return," or "high ESG" categories. Consequently, these shifts can change our conclusions about risk, pricing, and performance. The challenge is not merely technical. It is also interpretive. If a signal blends distinct components, it is unclear which component is responsible for an observed relation unless the construction is made explicit and the components are separated.

This dissertation studies three settings in which construction choices within these inputs are both pervasive and empirically consequential. The first setting examines how analysts' treatment of expected shares outstanding creates systematic differences between earnings per share (EPS) forecasts and firm-level earnings forecasts, and how these differences propagate into perceived earnings growth, forward valuation multiples, and implied expected-return measures. The second setting studies how the choice of earnings definition (Street versus GAAP) affects ICC estimation when forecast paths are mapped into implied discount rates. The third investigates how aggregation and ranking choices shape the informational content and portfolio implications of sustainability signals, benchmarking adverse-impact metrics and performance ranks against conventional ESG scores.

Across all three essays, the objective is not to argue that common conventions are generally wrong, but to treat construction rules as first-order objects of analysis: to make them explicit, quantify how much they matter for the measures researchers and practitioners use, and clarify interpretation in designs that rely on ranks, tails, and sorts. These settings are not chosen arbitrarily. Analyst forecasts are among the most widely used forward-looking inputs in capital markets research and practice. They shape valuation, serve as benchmarks for expectations, and enter applications that infer expected returns from prices (e.g., Francis and Soffer (1997); Gebhardt et al. (2001); Hou et al. (2012)). Similarly, sustainability measures have become influential in portfolio formation and disclosure regimes. However, existing evidence shows substantial disagreement among ESG scores from different providers (Chatterji et al. (2016); Berg et al. (2022)). This disagreement is not incidental, but reflects upstream construction choices that this dissertation treats as the object of study rather than background noise.

The first essay, *The EPS Growth Illusion: How Analysts' Share Count Forecasts Shape Valuation Perceptions*, studies how analyst EPS forecasts embed expectations about future shares outstanding and how these expectations shape implied earnings growth, valuation multiples, and expected return estimates. Although EPS forecasts are ubiquitous in practice and research, they are not simply firm-level earnings forecasts scaled by today's share count. Analysts forecast future earnings per share by combining two distinct expectations: expected firm-level earnings and expected changes in the share count driven by repurchases and issuance. Anticipated repurchases mechanically increase future EPS by reducing the denominator, thereby inflating implied earnings growth even if expected firm-level performance remains unchanged. From a benchmark valuation perspective, changes in the number of shares alone do not create value in the absence of changes in firm-level cash flows (Miller and Modigliani (1961)). Nevertheless, because many valuation metrics and empirical designs take per-share forecasts as the primary forward-looking input, the bundling of numerator and denominator expectations can alter perceived growth and relative valuation comparisons. Recent evidence on analyst share-count forecasts further motivates treating the share-count component as economically relevant rather than as a minor forecasting detail (Kaplan et al. (2024); Bradshaw et al. (2025)).

Using S&P 500 firms, where both EPS and firm-level forecasts are widely available and where repurchases and issuances are economically salient, the essay documents that forecast discrepancies are economically meaningful and increase with horizon. Median EPS forecasts exceed firm-level-based per-share forecasts by 0.77% at a one-year horizon and by more than 4% at a five-year horizon, with large right-tail cases. These level differences translate into systematically higher implied EPS growth rates relative to firm-level growth, on the order of one to two percentage points per year.

The divergence is systematic rather than random. Firms with strong expected repurchases tend to be more profitable and generate stronger free cash flows. However, they also face limited reinvestment opportunities. In contrast, expected issuers tend to have higher sales growth, but lower profitability and greater external financing needs. The processing choice has direct valuation consequences. Forward price–earnings ratios are mechanically lower for repurchasers (and higher for issuers) when computed with EPS forecasts, implying that buyback expectations can alter the cross-sectional ranking of apparent valuation even when underlying firm-level earnings expectations are held fixed. The essay further studies implications for implied costs of capital and, using the measurement-error-variance diagnostic framework of Lee et al. (2021), finds that ICCs based on firm-level earnings forecasts exhibit lower measurement-error variance and therefore provide a more reliable proxy for expected returns than EPS-based ICCs in this setting.

The second essay, *From Street to GAAP: Forecast Definitions in Implied Cost of Capital Estimation*, investigates how the choice of earnings definition – Street versus GAAP – shapes forecast properties and, consequently, forward-looking expected-return estimates. A large share of empirical research relies on analyst earnings forecasts as a summary of market expectations, but in many settings the earnings input is implicitly treated as a Street (non-GAAP) forecast. This convention reflects historical data availability and a literature arguing that non-GAAP measures exhibit stronger associations with abnormal returns in short-window pricing tests (Bradshaw and Sloan (2002); Bhattacharya et al. (2003)). More recent work emphasizes that GAAP–Street comparisons are not straightforward when expectations are measured with error or when the structural linkage between definitions is not respected (Bradshaw et al. (2018)). Street earnings equal GAAP earnings net of exclusions, so the Street–GAAP difference embeds an implicit expectation about excluded items. Empirical designs that compare Street and GAAP measures side by side without disciplining this decomposition can blur the distinction between the accounting-consistent component and the incremental adjustment component (Cohen et al. (2007); Christensen et al. (2025)).

This chapter revisits the Street-versus-GAAP question through the lens of ICC estimation, where forecast levels and implied growth across horizons jointly determine the inferred discount rate. Using matched Street and GAAP analyst earnings forecasts for S&P 500 firms from 2004 to 2024, I construct Street- and GAAP-based ICCs under five widely used models and a model-average composite, holding prices, book values, and implementation choices fixed so differences arise exclusively from forecast-definition inputs. The descriptive evidence establishes the economic mechanism: Street forecasts tend to be higher in levels, while GAAP forecasts frequently imply steeper growth across horizons. Since valuation models weigh short-run and long-run components differently, definition effects are model-dependent. Growth-sensitive models translate growth differences into particularly large ICC differences in the tails, with median magnitudes exceeding three percentage points for the most sensitive implementations. Even when Street- and GAAP-based ICCs are highly correlated, the choice of definition reshuffles firms across the ICC distribution in ways that matter for rank- and tail-based designs. This includes economically meaningful top-decile turnover even for the composite measure (e.g., 21.6%).

The final part of the essay examines the validity of alternative ICC constructions as proxies for expected returns, as well as their relationship to subsequent realized returns. Following the methodology outlined in Lee et al. (2021), measurement-error diagnostics indicate broadly similar aggregate reliability for Street- and GAAP-based ICCs. Motivated by the decomposition of Street forecasts into a GAAP-consistent component and an adjustment component, I show that both GAAP-based ICC and the difference in ICCs are positively related to future returns at short horizons, with the return relevance of the GAAP-based ICC strongest when forecast divergence is high and in settings associated with greater informational frictions, such as smaller firms and intangible-intensive industries. Overall, the findings imply that ICC studies cannot treat the choice between Street and GAAP forecasts as a minor implementation variation because it can alter cross-sectional ordering and change conclusions based on returns.

The third essay, *Beyond ESG Scores: Risk and Return Insights from SFDR Metrics*, shifts attention to sustainability information and the way it is translated into investment-relevant signals. ESG considerations have become central to investment practice and regulation, yet a persistent obstacle to cumulative evidence is measurement. Conventional ESG scores often bundle heterogeneous dimensions into single composites, and scores diverge substantially across providers due to differences in indicator selection, weighting schemes, and aggregation methods (e.g., Chatterji et al. (2016); Berg et al. (2022)). Such divergence can obscure the channels through which sustainability characteristics map into financial risk and can inject noise into portfolio tests.

To address this issue, the essay adopts a measure based on Principal Adverse Impact (PAI) indicators under the EU Sustainable Finance Disclosure Regulation (SFDR). These indicators provide a standardized, regulation-based framework that targets negative externalities plausibly linked to regulatory, reputational, and financial risk (European Commission (2022)). The chapter further emphasizes an underappreciated processing margin in ESG applications: how information is normalized and ranked. Building on evidence that percentile-based methods can generate distorted and persistent rankings, the essay compares conventional percentage-rank metrics with performance-rank (PR) metrics that more directly reflect relative sustainability performance (Benuzzi et al. (2025)).

Using London Stock Exchange Group (LSEG) data, the chapter constructs SFDR-based metrics under both ranking schemes and benchmarks their portfolio implications against strategies formed on standard ESG metrics. The results show that SFDR-based PR measures dominate common alternatives in the studied portfolio tests: they generate return spreads of up to 0.77 percentage points and higher Sharpe ratios in best-in-class strategies. More generally, performance-rank normalization improves outcomes relative to percentile ranking, increasing return spreads by up to 0.34 percentage points and Sharpe ratios by up to 0.19. Overall, the findings show that measurement and processing choices are economically meaningful for sustainable investing, and that combining SFDR-based indicators with performance-rank normalization yields the most informative signal in the settings studied.

The dissertation's main contribution is to treat the construction of widely used market signals as part of the economics rather than a technical footnote. Across three settings, it exploits designs in which key inputs can be held fixed while one construction margin changes, yielding clean evidence on how measurement choices propagate into growth perceptions, valuation ratios, implied discount rates, and portfolio classifications. For academic research, this perspective adds value by clarifying what is – and is not – identified when studies sort firms, assign tails, or run return tests on constructed measures, and by motivating specifications that either separate mechanically distinct components or transparently report sensitivity to definitional choices. More broadly, it reframes common "inputs" (earnings forecasts, ICCs, ESG scores) as outcomes of measurement architecture, which helps reconcile why seemingly similar implementations can produce different rank orderings and inferences. For practitioners, the primary implication is one of measurement awareness: forward multiples, growth narratives, ICC estimates, and ESG rankings can shift as much due to construction conventions as to changes in underlying fundamentals. Metric definitions, scaling choices, and scoring methods should therefore be treated as first-order considerations when comparing firms or constructing portfolios.

To conclude the introduction, I briefly summarize my contributions to the three essays in this thesis. The first essay originated from joint discussions with my co-author. We jointly developed the research design and conducted most of the empirical analyses. I prepared the initial draft, and the manuscript was revised collaboratively over multiple rounds. The second essay is single-authored. I developed the research idea, implemented the empirical analysis, and wrote the paper. The third essay builds on an initial research idea proposed by me, which was subsequently refined jointly through continued discussions. My co-author conducted the majority of the empirical analysis, incorporating feedback from our iterative exchanges. I wrote the initial draft, and the paper was revised jointly.

Chapter 2

The EPS Growth Illusion: How Analysts' Share Count Forecasts Shape Valuation Perceptions*

2.1 Introduction

Earnings per share (EPS) forecasts issued by equity analysts are a central input in investment practice and empirical research. Investors rely on EPS forecasts to assess firm performance, form growth expectations, and evaluate valuation metrics such as forward price–earnings ratios. In academic research, analyst EPS forecasts are widely used as proxies for market expectations and as key inputs in estimating implied costs of capital (ICC). Owing to their availability and perceived interpretability, EPS forecasts are commonly treated as scaled versions of firm-level earnings forecasts.

However, EPS forecasts are not merely firm-level (FL) earnings divided by the current number of shares outstanding. In practice, analysts construct EPS forecasts using expectations about future shares outstanding, thereby embedding assumptions about share repurchases and equity issuances. As a result, EPS forecasts combine two distinct components: expectations about future firm-level earnings and expectations about future changes in the firm's share count. Consequently, an EPS forecast can differ mechanically from a firm-level earnings forecast scaled by today's shares outstanding, even when

*This paper is based on Hess and Zährl (2025). We thank Tobias Bauckloh, Mario Hendriock, Alexander Kempf, Maximilian Müller, Harm Schütt and Tom Zimmermann for insightful discussions. We are also grateful for helpful comments and suggestions from participants of research workshops at the Centre for Financial Research (CFR) and the German Bundesverband Investment und Asset Management (BVI).

analysts share identical views about underlying economic performance.

This distinction matters because it directly affects how earnings growth and valuation inputs are constructed. Anticipated share repurchases mechanically increase future EPS by reducing the denominator, thereby inflating implied earnings growth without any change in firm-level earnings. Conversely, expected equity issuances mechanically depress EPS growth. Under standard valuation frameworks, changes in share count alone should not affect firm value when firm-level earnings are unchanged (e.g., Miller and Modigliani (1961)). Nevertheless, when valuation metrics rely on EPS forecasts, embedded share-count expectations can distort perceived growth rates, valuation multiples, and expected returns, even though the firm's underlying profitability remains the same.

To illustrate this mechanism, consider Apple as of December 2024. Analysts forecast Apple's EPS to be \$7.40 in 2025. In contrast, dividing analysts' projected firm-level earnings by the current number of shares outstanding yields an FL-based forecast of only \$7.17.[1] Growth expectations also differ meaningfully: using analysts' EPS forecasts implies an average expected EPS growth rate of 10.5% p.a. over the subsequent three years, whereas growth based on FL forecasts is only 8.8% per year.[2] This divergence arises because analysts expect Apple's share count to decline by approximately 2.3% per year over this period. Thus, part of the perceived earnings growth reflects expected buybacks rather than improvements in operating performance. This example illustrates that EPS forecasts can systematically differ from firm-level earnings forecasts because they embed assumptions about future share counts.

Despite the widespread use of analyst earnings forecasts, relatively little is known about how analysts form expectations about future shares outstanding and how these expectations affect the interpretation of EPS. To the best of our knowledge, only two recent studies directly examine analyst share forecasts. Kaplan et al. (2024) study the accuracy of share forecasts and show that EPS forecasts incorporating expected shares can outperform alternatives based on realized shares. Bradshaw et al. (2025) document that explicit share forecasts improve EPS accuracy and influence managerial behavior. However, these studies focus on forecast accuracy and behavioral implications, leaving unexplored

[1] The FL-based forecast is calculated by dividing the firm-level earnings forecast by the most recently released weighted average number of shares to calculate diluted EPS: $\frac{E_t[E_{t+1}]}{\text{cshfd}_t} = \frac{11{,}453}{15{,}408}$.

[2] The FL-based growth is calculated as $\sqrt[3]{\frac{E_t[E_{t+3}]}{E_t[E_t]}} - 1 = \sqrt[3]{\frac{133{,}773}{103{,}982}} - 1 = 8.8\%$. The EPS-based growth is calculated as $\sqrt[3]{\frac{E_t[\text{EPS}_{t+3}]}{E_t[\text{EPS}_t]}} - 1 = \sqrt[3]{\frac{9.10}{6.75}} - 1 = 10.5\%$.

how embedded share-count expectations affect valuation metrics and expected return estimates used in both research and practice.

This paper addresses this gap by asking four related questions. First, how large are the discrepancies between EPS- and firm-level (FL) earnings forecasts? Second, how do these discrepancies affect forward price-earnings ratios, a frequently applied valuation metric? Third, what are the implications for ICC estimates, and are ICCs based on FL forecasts more reliable than those based on EPS? Fourth, which firm characteristics predict expected and realized repurchases and issuances, and how accurate are analysts' share-count forecasts?

We focus on S&P 500 firms, for which both EPS and FL earnings forecasts are largely available. These firms frequently engage in share repurchases and equity issuances, making them a natural setting to study how analysts incorporate share-count expectations into earnings forecasts. Consistent with this institutional context, we document that for 10% of S&P 500 firms, analysts expect annual share repurchases to exceed 4% at any given point in time over the past decade.

Our empirical results show that differences between EPS- and FL-based earnings forecasts are economically meaningful and grow with the forecast horizon. Among S&P 500 firms, median EPS forecasts exceed firm-level forecasts by 0.77% at a one-year horizon and by more than 4% at a five-year horizon, with extreme cases exceeding 29%. These differences translate into systematically higher implied earnings growth rates when forecasts are expressed on a per-share basis, with EPS-based growth exceeding firm-level growth by roughly one to two percentage points per year. Moreover, we document that these effects are not random: firms expected to repurchase shares tend to be more profitable, generate stronger free cash flows, and face limited reinvestment opportunities, whereas firms expected to issue equity exhibit higher sales growth but lower profitability and greater external financing needs.

These differences in earnings forecasts have direct implications for valuation. We find that forward price–earnings ratios are highly sensitive to the choice of earnings concept. Firms expected to repurchase shares display artificially lower forward P/E ratios due to mechanically inflated EPS forecasts, while firms expected to issue shares exhibit higher P/E ratios due to earnings dilution. For example, in the top decile of firms with the strongest expected buybacks, forward P/E ratios are more than 7% lower when using EPS forecasts rather than firm-level earnings.

Analyst earnings forecasts are also a key input in estimating the implied cost of capital, a forward-looking measure of expected returns inferred from current stock prices and forecasts of future earnings rather than from historical returns. Because ICC estimates rely directly on forecasted earnings, their accuracy depends critically on the quality and structure of the underlying forecast inputs. While prior research examines biases in analyst forecasts and alternative ICC estimation techniques (e.g., Gode and Mohanram (2003); Easton and Monahan (2005); Hou et al. (2012)), little attention has been paid to how the use of firm-level versus per-share earnings forecasts affects ICC estimates.

We document that differences in earnings forecasts also distort ICC estimates. EPS-based ICCs are systematically higher for firms expected to repurchase shares and lower for firms expected to issue shares. Discrepancies between EPS- and FL-based ICCs can be substantial. For example, among firms in the top decile of expected buybacks, EPS-based ICCs exceed firm-level-based ICCs by more than 3 percentage points at the median for the most sensitive models. Moreover, using the evaluation framework of Lee et al. (2021), we find that ICCs based on firm-level earnings exhibit lower measurement error variance than EPS-based ICCs, indicating that firm-level earnings provide a more reliable basis for estimating expected returns.

This paper makes several contributions. First, we document and quantify systematic discrepancies between EPS and firm-level earnings forecasts that arise from analysts' expectations about future share counts. Second, we provide the first systematic evidence that these discrepancies materially distort forward-looking valuation metrics, including forward price–earnings ratios and ICCs. Third, our findings highlight important methodological implications for research using analyst forecasts. Studies that rely on EPS forecasts – particularly those evaluating forecast accuracy or estimating expected returns – may inadvertently conflate earnings expectations with share-count assumptions. Our results suggest that firm-level earnings forecasts offer a cleaner and more reliable input in these settings.

More broadly, our findings underscore the importance of transparency when selecting and interpreting earnings forecasts in both empirical research and investment practice. While incorporating expected share-count changes may improve forecast accuracy, doing so can materially affect growth perceptions, valuation ratios, and expected return estimates. Recognizing this distinction is essential for accurately interpreting analyst forecasts and the valuation metrics derived from them.

The remainder of the paper is organized as follows. Section 2.2 discusses the theoretical background and related literature. Section 2.3 describes the data and variable construction. Section 2.4 presents the empirical results, and Section 2.5 concludes.

2.2 Theoretical background

A substantial body of literature has examined the role, accuracy, and usefulness of analyst earnings forecasts in financial markets. Analyst forecasts are a critical input for investors, who use them in valuation models to estimate firms' intrinsic value and inform trading and asset allocation decisions (e.g., Francis and Soffer (1997)). Analysts are commonly viewed as informed intermediaries who synthesize public and private information, and their forecasts often move markets, particularly when associated with earnings revisions or surprises (e.g., Livnat and Mendenhall (2006)). Beyond investment decisions, analyst forecasts are also used in corporate decision-making, such as guiding investor relations strategies and managing earnings expectations (e.g., Graham et al. (2005)). In academic research, analyst forecasts serve as proxies for market expectations, benchmarks for testing earnings management, and key inputs in models estimating the implied cost of capital (e.g., Gebhardt et al. (2001)). Despite their widespread use, analyst forecasts are subject to systematic biases such as optimism, herding, and conflicts of interest (e.g., Brown (2001); Hong and Kubik (2003)), and their accuracy varies with analyst experience, firm characteristics, and macroeconomic conditions (e.g., Mikhail et al. (1997)). In addition, forecast coverage is uneven: smaller, less visible, or financially distressed firms are often underrepresented, and long-term forecasts are frequently unavailable, particularly in earlier periods (e.g., Diether et al. (2002)). These limitations result in restricted cross-sectional and time-series coverage.

Nevertheless, due to their market relevance, analyst forecasts remain a foundational tool in both financial practice and empirical research. Importantly, analysts provide forecasts for a range of financial metrics, and the availability of these forecasts in commercial databases has evolved over time. EPS forecasts have been available in the IBES database since 1978, whereas firm-level earnings forecasts only became available in 2002. Forecasts for shares outstanding have only been available since 2013 in commercial databases and are still not included in the widely used *LSEG IBES Academic* data package provided through WRDS. Access to such forecasts through WRDS has only recently become available via the newly introduced *LSEG IBES KPI* package. This limited and relatively recent availability likely explains why analyst forecasts of future shares outstanding—and their role in shaping EPS forecasts—have received relatively little attention.

To date, only two studies directly examine analyst forecasts of future shares outstanding. Kaplan et al. (2024) shows that while share forecasts are often inaccurate, EPS forecasts incorporating expected shares can outperform alternatives based on realized shares. The study of Bradshaw et al. (2025) documents that explicit share forecasts improve EPS accuracy and influence managerial behavior, such as reducing the likelihood of accretive share repurchases. Although these studies demonstrate the informational value of share-count forecasts, they do not examine how such forecasts affect the interpretation of EPS or their implications for valuation and expected return estimation.

At first glance, EPS and firm-level earnings forecasts may appear interchangeable, as one might assume they can be easily converted by multiplying or dividing by the current number of shares outstanding. However, a crucial distinction is that per-share forecasts are based on forecasted future shares outstanding rather than the current share count. As a result, EPS forecasts embed assumptions about future share repurchases or equity issuances, whereas firm-level earnings forecasts do not. This difference implies that earnings levels and growth rates derived from per-share forecasts can diverge substantially from those based on firm-level earnings, even when analysts hold identical expectations about a firm's underlying economic performance. Anticipated share repurchases mechanically increase future EPS by reducing the denominator, thereby inflating implied growth rates without any change in firm-level earnings. Conversely, expected equity issuances mechanically depress EPS growth. While Bradshaw et al. (2018) highlight the value of newly available forecast data in advancing accounting research, the availability of share-count forecasts introduces an additional dimension that affects the interpretation of EPS-based measures.

A large literature evaluates the implied cost of capital as a proxy for expected returns. Studies in this area typically examine whether ICC estimates are informative about future realized returns (e.g., Gebhardt et al. (2001); Gode and Mohanram (2003); Easton and Monahan (2005); Hou et al. (2012); Li et al. (2013); Li and Mohanram (2014); Callen and Lyle (2020)). The underlying premise is that a reliable expected return proxy should exhibit a positive association with subsequent realized returns. However, evidence in this literature remains mixed. Some studies document a positive relation between ICC and realized returns (e.g., Gode and Mohanram (2003); Hou et al. (2012); Li et al. (2013); Li and Mohanram (2014); Callen and Lyle (2020)), while others find no significant relation (e.g., Gebhardt et al. (2001); Easton and Monahan (2005); Guay et al. (2011)) or even a negative relation (Kang and Sadka (2017)). A key challenge in this literature concerns the quality of the earnings forecasts used as inputs in ICC estimation, as forecast bias directly affects inferred expected returns. Notably, Hou et al. (2012) show that less biased earnings forecasts lead to more accurate return predictions. More recently, Lee et al. (2021) propose a framework for evaluating expected return proxies based on their measurement error properties, highlighting the importance of forecast inputs in assessing ICC reliability.

Related work has also examined the accuracy of earnings estimates and their implications for cost of equity calculations (e.g., Abarbanell and Bushee (1997); Easton and Monahan (2005)). Analyst forecasts are often found to be overly optimistic, leading to upwardly biased cost of equity estimates when used in ICC models (e.g., Easton and Sommers (2007)). To address these concerns, cross-sectional models have been developed to estimate future earnings and serve as alternatives to analyst forecasts in ICC estimation (e.g., Hou et al. (2012); Li and Mohanram (2014)). However, even in these comparative frameworks, analyst earnings forecasts—typically in the form of EPS—are used as benchmarks without accounting for embedded assumptions about future share counts. This omission is important because expected changes in shares outstanding due to repurchases or issuances can mechanically distort EPS-based forecasts relative to firm-level earnings. In addition, ICC calculations often scale firm-level earnings using the current number of shares outstanding (see Hou et al. (2012); Li and Mohanram (2014)), creating a methodological inconsistency when combined with EPS-based forecasts that reflect expected changes in share counts. Such inconsistencies can lead to systematic estimation errors and complicate the interpretation of ICC estimates.

Finally, this study relates to the literature on share repurchases and equity issuances, which provides important context for understanding analysts' share-count expectations. Corporate repurchases and issuances play a central role in shaping EPS, capital structure, and investor perceptions. While repurchases are often interpreted as signals of undervaluation or excess cash, early studies document positive market reactions and long-run abnormal returns following buyback announcements (e.g., Vermaelen (1981); Ikenberry et al. (1995)). Equity issuances, by contrast, are frequently associated with overvaluation and subsequent underperformance, consistent with market-timing behavior by managers (e.g., Myers and Majluf (1984); Loughran and Ritter (1995); Baker and Wurgler (2002)). At the same time, growing evidence suggests that repurchases may be used strategically to manage reported performance rather than to enhance firm value. For example, Hribar et al. (2006) find that firms repurchase shares to avoid EPS misses, and Bens et al. (2003) document increased buyback activity related to executive option granting. More broadly, Almeida et al. (2016) show that repurchases can crowd out investment, raising concerns about the long-term consequences of payout policies driven by short-term performance targets.

Overall, the literature suggests that earnings forecasts are inherently sensitive to assumptions about future share counts, particularly for firms that actively engage in repurchase or issuance programs. Yet relatively few studies have examined how such assumptions are incorporated into analyst forecasts or how they affect the interpretation and use of earnings-based valuation metrics.

2.3 Data

Our sample consists of S&P 500 firms covered in the LSEG database over the period from 2004 through 2024. We focus on large, widely followed firms because both per-share and firm-level earnings forecasts are more consistently available for this group than for smaller firms. The sample comprises 125,750 firm-month observations (with 121,809 EPS forecasts and 121,764 firm-level forecasts for FY1) from 958 distinct firms. All financial statement items and analyst forecasts are denominated in U.S. dollars. We exclude firms with multiple share classes to ensure consistency in the measurement of shares outstanding and earnings per share.

We obtain analyst earnings forecasts, firm-level earnings forecasts, and forecasts of shares outstanding directly from LSEG. To assess data consistency, we cross-check key forecast variables and financial statement items against Compustat and IBES datasets available through WRDS. This comparison reveals no material discrepancies, supporting the reliability of the underlying data used in our analysis.

Table 2.1 reports the annual availability of LSEG forecast variables from 2004 to 2024 for S&P 500 firms. While EPS and firm-level earnings forecasts are available throughout the sample period, forecasts of shares outstanding are only available from 2013 onward.

To maintain comparability between per-share and firm-level earnings measures over the full sample period, we scale firm-level earnings by the diluted weighted average number of shares outstanding when constructing per-share equivalents, rather than retroactively adjusting EPS forecasts to firm-level figures. This approach ensures consistency in earnings measurement while preserving the original structure of analysts' forecasts.

Table 2.1: Availability of LSEG forecast variables

Year	#Shares$_{F1}$	EPS$_{F1}$	FL$_{F1}$
2004	0	5,660	5,615
2005	0	5,649	5,609
2006	0	5,626	5,647
2007	0	5,643	5,658
2008	0	5,646	5,641
2009	0	5,662	5,646
2010	0	5,681	5,698
2011	0	5,681	5,684
2012	0	5,679	5,676
2013	2,788	5,704	5,714
2014	5,737	5,736	5,737
2015	5,814	5,812	5,814
2016	5,841	5,841	5,840
2017	5,878	5,878	5,876
2018	5,914	5,913	5,915
2019	5,945	5,952	5,947
2020	5,954	5,967	5,967
2021	5,981	5,993	5,993
2022	6,019	6,018	6,020
2023	6,035	6,036	6,034
2024	6,030	6,032	6,033
Total	67,936	121,809	121,764

This table reports for S&P500 (SPX) firms the number of LSEG observations per year for the variables: #Shares$_{F1}$, EPS$_{F1}$, and FL$_{F1}$. The variable #Shares$_{F1}$ represents the median analysts' forecast for the one-year-ahead number of outstanding shares. EPS$_{F1}$ denotes the median one-year-ahead forecast for earnings per share (EPS). FL$_{F1}$ represents the firm-level earnings forecast provided by analysts for one year ahead. The sample covers the period 2004–2024.

2.4 Results

This section presents four sets of results. First, we document systematic differences between EPS and FL earnings forecasts and show that these differences increase with the forecast horizon. Second, we examine the implications for forward price-earnings ratios (FPE) and show that valuation multiples depend mechanically on whether earnings forecasts embed expected share count changes. Third, we quantify the impact on ICC estimates and evaluate the reliability of EPS- versus FL-based ICCs using the framework of Lee et al. (2021). Finally, we analyze firm characteristics associated with expected and realized share repurchases and issuances and study the accuracy of analysts' share-count forecasts.

2.4.1 Discrepancies in earnings forecasts

We begin by examining analysts' forecasts of changes in shares outstanding over one- to five-year horizons. Table 2.2 reports the cross-sectional distribution of these forecasts from 2013 to 2024. For each firm-year, forecasted share growth is defined as the expected cumulative percentage change in shares outstanding over the relevant horizon (FY1 to FY5), measured as consensus analyst projections of future shares outstanding relative to the most recent fiscal year-end share count.

Panel A shows that expected share growth is slightly negative at short horizons and becomes more negative as the forecast horizon increases. The median projected share growth is -0.76% for one-year-ahead forecasts (FY1) and declines to -3.42% by FY5. Similarly, the mean declines from 0.08% to -3.71%. These values indicate that, on average, analysts expect net share repurchases rather than issuances. The tails are economically large and become more pronounced at longer horizons. By FY1, the 10th percentile implies a cumulative share reduction of 5.16% over the next year, while the 90th percentile implies a cumulative share increase of 2.99%. By FY5, these magnitudes widen to 18.63% and 7.41%, respectively. Thus, analysts expect substantial share repurchases and issuances in both tails, with expected net repurchases in the lower tail roughly twice as large as expected net issuances in the upper tail.

Panel B provides a complementary view by sorting firms into deciles of expected share growth. Firms in the lowest decile (D1), expected to repurchase the largest proportion of shares, show median growth forecasts declining from -6.96% in FY1 to -24.39% in FY5. Firms in the highest decile (D10), expected to issue the largest proportion of shares, exhibit growth forecasts increasing from 8.54% in FY1 to 15.29% in FY5. The spread between D1 and D10 widens markedly over the forecast horizon and reaches 40 percentage points by FY5, underscoring the increasing economic importance of expected share-count adjustments at longer horizons.

Figure 2.1 plots annual cross-sectional percentiles (10th, 25th, 50th, 75th, and 90th) of growth implied in analysts' one-year-ahead forecasts of shares outstanding, defined

Table 2.2: Forecasted number of shares growth

Panel A: Summary Statistics					
Statistic	$\Delta\text{Shares}_{FY1}$	$\Delta\text{Shares}_{FY2}$	$\Delta\text{Shares}_{FY3}$	$\Delta\text{Shares}_{FY4}$	$\Delta\text{Shares}_{FY5}$
N	66,155	66,146	64,246	48,112	37,861
Mean	0.08%	-0.74%	-1.87%	-2.74%	-3.71%
Median	-0.76%	-1.50%	-2.48%	-2.89%	-3.42%
P10	-5.16%	-8.74%	-12.17%	-15.60%	-18.63%
P90	2.99%	4.27%	5.27%	6.49%	7.41%
Panel B: Decile Medians					
Decile	$\Delta\text{Shares}_{FY1}$	$\Delta\text{Shares}_{FY2}$	$\Delta\text{Shares}_{FY3}$	$\Delta\text{Shares}_{FY4}$	$\Delta\text{Shares}_{FY5}$
D1	-6.96%	-11.39%	-15.53%	-20.29%	-24.39%
D2	-4.09%	-7.10%	-10.04%	-12.61%	-15.11%
D3	-2.77%	-4.81%	-6.99%	-8.94%	-10.83%
D4	-1.83%	-3.29%	-4.95%	-6.24%	-7.50%
D5	-1.09%	-2.07%	-3.25%	-3.97%	-4.68%
D6	-0.43%	-0.96%	-1.64%	-1.79%	-2.19%
D7	0.01%	-0.03%	-0.22%	-0.09%	-0.17%
D8	0.42%	0.53%	0.58%	0.91%	0.97%
D9	1.38%	2.04%	2.67%	3.74%	4.20%
D10	8.54%	10.85%	12.05%	13.92%	15.29%

This table reports the cross-sectional distribution of the forecasted number of shares growth across forecast horizons of one to five fiscal years ahead ($FY1$–$FY5$). Panel A presents overall summary statistics, including the number of observations (N), mean, median, and the 10th and 90th percentiles of the forecasted growth rates. Panel B reports the median forecasted growth rate within each decile of the full sample distribution. Forecasted share growth for each firm-year and forecast horizon i is calculated as: $\Delta\text{Shares}_{FYi} = \left(\frac{\text{Forecasted Shares}_{t+i}}{\text{Actual Shares}_t} - 1\right) \times 100$, where $\text{Forecasted Shares}_{t+i}$ is the analyst consensus forecast for the number of shares outstanding at horizon i, and Actual Shares_t is the current (fiscal year t) number of shares outstanding. All values are expressed in percentage terms. The sample covers the period 2013–2024.

as the consensus forecast of next fiscal year-end shares outstanding relative to the most recent fiscal year-end share count. The median is below zero in most years, consistent with expected net repurchases, and moves closer to zero around the onset of COVID-19, suggesting reduced buyback activity and/or increased issuances. While the level of the percentiles varies over time, the overall pattern is persistent: in most years, more than 10% of firms are expected to repurchase more than 4% of shares within one year, while another 10% are expected to issue more than 2%. These tails remain present throughout the sample, indicating that the summary statistics are not driven by a small number of outlier years. These tail firms are particularly relevant for assessing how share-count expectations affect per-share earnings forecasts.

Next, we examine how these expected share-count changes translate into differences between EPS and FL earnings forecasts. Table 2.3 reports descriptive statistics for analysts' EPS and FL forecasts over horizons from one to five years ahead (FY1–FY5), together

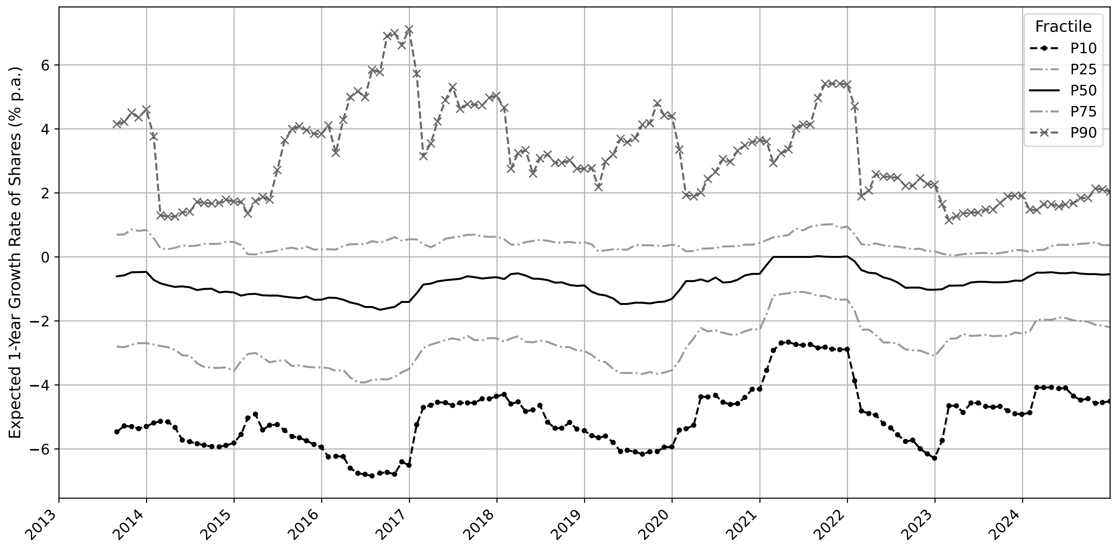


Figure 2.1: Forecasted growth in outstanding shares

This figure displays the annual cross-sectional 10%, 25%, 50%, 75%, 90% fractiles of the forecasted 1-year-ahead growth rate in shares outstanding, i.e., analysts' median forecast of number of shares outstanding at next fiscal year end date (Refinitiv item: $NumberofSharesOutstandingMedian$) relative to the actual outstanding number of shares at the forecast date ($NumberofSharesOutstandingActual$) for 2013–2024. Values below zero show expected net repurchases while values above zero represent net issuances.

with the dollar and percentage differences between the two forecast types.

Panel A shows that both EPS and FL forecasts increase with the horizon. The mean rises from approximately \$4.37 for EPS and \$4.34 for FL in FY1 to \$6.93 (EPS) and \$6.60 (FL) in FY5, and the dispersion increases with horizon length. Across all horizons, EPS forecasts are consistently higher than FL forecasts, and the gap widens with the horizon, consistent with the cumulative effect of expected net share repurchases.

Panels B and C relate the EPS–FL wedge to expected share-count changes. Firms are sorted into deciles based on forecasted share growth (decile 1 denotes the largest expected reductions in shares, and decile 10 the largest expected increases). The results show a clear monotonic pattern. In decile 1, the median EPS forecast exceeds the FL forecast by 7.07% (\$0.35) in FY1, increasing to 29.11% (\$1.78) in FY5. In contrast, in decile 10, EPS forecasts are below FL forecasts by −6.84% (−\$0.18) in FY1 and −11.97% (−\$0.58) in FY5. These patterns indicate that expected changes in shares outstanding are a primary driver of the divergence between EPS and FL forecasts, and that the magnitude of the wedge increases with the forecast horizon.

Table 2.4 reports median growth rates implied by EPS and FL forecasts. Panel A shows growth across five consecutive annual periods. While median growth declines

Table 2.3: Earnings forecasts and differences by share growth deciles

Panel A: Summary Statistics of Earnings Forecasts

Statistic	Earnings Type	FY1	FY2	FY3	FY4	FY5
P10	EPS	0.8300	1.0800	1.2900	1.4500	1.6000
	FL	0.8428	1.0868	1.2893	1.4397	1.5519
Median	EPS	3.0500	3.4100	3.8150	4.2460	4.6800
	FL	3.0487	3.3948	3.7544	4.1257	4.4942
Mean	EPS	4.3687	4.9248	5.4968	6.1450	6.9271
	FL	4.3436	4.8514	5.3421	5.8991	6.6041
P90	EPS	9.1300	10.0400	11.1100	12.4000	13.9079
	FL	9.0395	9.9001	10.8013	11.9269	13.3678
N		116,923	116,923	116,923	116,923	116,923

Panel B: Percent Differences by Share Growth Decile

Decile	FY1	FY2	FY3	FY4	FY5
D1	7.07%	12.46%	17.62%	23.66%	29.11%
D2	4.21%	7.61%	10.97%	14.06%	17.77%
D3	2.85%	5.12%	7.48%	9.71%	12.05%
D4	1.90%	3.50%	5.28%	6.75%	8.13%
D5	1.15%	2.26%	3.55%	4.24%	4.94%
D6	0.48%	1.17%	1.98%	1.98%	2.55%
D7	-0.02%	0.11%	0.42%	0.08%	0.13%
D8	-0.38%	-0.48%	-0.49%	-0.92%	-0.93%
D9	-1.27%	-1.96%	-2.47%	-3.18%	-3.82%
D10	-6.84%	-8.96%	-10.11%	-11.48%	-11.97%
N	65,165	65,154	63,292	47,400	37,293

Panel C: Dollar Differences by Share Growth Decile

Decile	FY1	FY2	FY3	FY4	FY5
D1	0.3480	0.6463	0.9720	1.4416	1.7840
D2	0.1959	0.3810	0.5963	0.8276	1.1703
D3	0.1358	0.2624	0.4165	0.6011	0.8286
D4	0.0859	0.1726	0.2861	0.3820	0.5435
D5	0.0480	0.1129	0.1933	0.2510	0.3514
D6	0.0190	0.0517	0.1003	0.1024	0.1529
D7	-0.0006	0.0033	0.0165	0.0037	0.0065
D8	-0.0125	-0.0165	-0.0180	-0.0393	-0.0523
D9	-0.0365	-0.0692	-0.0939	-0.1344	-0.1878
D10	-0.1794	-0.2811	-0.3611	-0.4862	-0.5780
N	65,165	65,154	63,292	47,400	37,293

Panel A reports the mean, median, 10th percentile (P10), and 90th percentile (P90) for earnings forecasts based on EPS and FL estimates over forecast horizons of one to five years ahead (FY1–FY5) for the period 2004–2024. Panels B and C report median forecast differences between the consensus EPS forecast and the consensus FL forecast across deciles of the forecasted number-of-shares growth distribution. For each fiscal-year-ahead horizon ($FY1$–$FY5$), firms are sorted into deciles based on their forecasted number-of-shares growth, computed as $\Delta\text{Shares}_{FYi} = \left(\frac{\text{Forecasted Shares}_{t+i}}{\text{Actual Shares}_t} - 1\right) \times 100$. Panel B presents percent differences, defined as $\%Diff = \frac{\text{EPS}_{FYi} - \text{FL}_{FYi}}{\text{FL}_{FYi}} \times 100$. Panel C reports dollar differences, defined as $Diff = \text{EPS}_{FYi} - \text{FL}_{FYi}$. Across both panels, values represent medians within each decile; decile 1 contains firms with the most negative expected share growth (buybacks) and decile 10 contains firms with the largest expected share increases (issuances). The sample for Panels B and C includes all available analyst consensus forecasts from 2013–2024.

modestly as the horizon extends, EPS-based growth exceeds FL-based growth in each period by roughly one to two percentage points. For example, the median growth from year one to year two is 11.11% using EPS forecasts versus 9.78% using FL forecasts, declining to 10.02% versus 8.49% by year five. Panel B reports average growth rates over horizons from one to five years and yields a similar conclusion: over five years, EPS forecasts imply average growth of 10.42% p.a., whereas FL forecasts imply 8.95% p.a. Taken together, Tables 2.3 and 2.4 show that EPS forecasts imply systematically higher earnings levels and growth rates than FL forecasts, consistent with analysts embedding expectations of net share repurchases into per-share forecasts.

Table 2.4: Median growth rates implied in earnings forecasts

Panel A: Year-to-Year Growth Rates					
Earnings Type	L0→F1	F1→F2	F2→F3	F3→F4	F4→F5
EPS	9.46%	11.11%	11.08%	10.73%	10.02%
FL	8.17%	9.78%	9.38%	9.20%	8.49%
Panel B: Average Annual Growth Rates					
Earnings Type	L0→F1	L0→F2	L0→F3	L0→F4	L0→F5
EPS	9.46%	10.50%	10.56%	10.55%	10.42%
FL	8.17%	9.19%	9.15%	9.08%	8.95%

This table reports the median growth rates implied by earnings forecasts based on analyst EPS estimates and firm-level (FL) forecasts. Panel A presents median year-over-year forecasted earnings growth between consecutive forecast horizons (L0→F1 through F4→F5). Panel B shows median cumulative average annual growth rates over increasingly longer forecast horizons (L0→F1 to L0→F5). All growth rates are expressed in percentage terms. The sample period covers fiscal years 2004–2024.

2.4.2 Impact on forward price–earnings ratio

Differences between EPS and FL forecasts affect forward price–earnings (FPE) ratios because EPS-based valuation multiples embed expected changes in shares outstanding.

The forward P/E ratio based on EPS is

$$\text{FPE}_{EPS} = \frac{P_0}{EPS_1}. \tag{2.1}$$

To isolate the role of expected share-count changes, we rewrite price and earnings in firm-level terms:

$$P_0 = \frac{MarketCap_0}{Shares_0}, \qquad EPS_1 = \frac{FL_1}{Shares_1}.$$

Substituting these identities into (2.1) yields

$$\mathrm{FPE}_{EPS} = \frac{\frac{MarketCap_0}{Shares_0}}{\frac{FL_1}{Shares_1}} = \frac{MarketCap_0}{FL_1} \cdot \frac{Shares_1}{Shares_0}.$$

The FL-based forward P/E ratio is defined as

$$\mathrm{FPE}_{FL} = \frac{MarketCap_0}{FL_1}.$$

Therefore,

$$\mathrm{FPE}_{EPS} = \mathrm{FPE}_{FL} \cdot \frac{Shares_1}{Shares_0}. \tag{2.2}$$

Equation (2.2) shows that the wedge between FPE_{EPS} and FPE_{FL} is mechanically given by $\frac{Shares_1}{Shares_0}$. Expected net repurchases imply $\frac{Shares_1}{Shares_0} < 1$ and therefore lower FPE_{EPS} relative to FPE_{FL}; expected net issuances imply the opposite. Consequently, cross-sectional comparisons based on EPS-based FPEs can reflect anticipated share-count changes rather than differences in firm-level earnings.

Table 2.5 reports FPE ratios based on EPS and FL forecasts. Panel A shows that the mean FPE ratios are 20.73 for FPE_{EPS} and 25.15 for FPE_{FL}, while medians are similar (16.93 versus 17.05), consistent with a right-skewed distribution. The higher mean for FPE_{FL} reflects that FL forecasts are, on average, lower than EPS forecasts, consistent with the negative median expected share growth documented in subsection 2.4.1.

Panel B examines heterogeneity by sorting firms into deciles based on one-year-ahead expected share growth, $\Delta\mathrm{Shares}_{FY1}$. Differences between EPS- and FL-based FPE ratios are concentrated in the tails. Expected issuances raise EPS-based FPE ratios mechanically (D10: 21.46 versus 19.04; 7.10%), whereas expected repurchases mechanically compress them (D1: 13.63 versus 14.69; −6.62%), causing issuing firms to appear more expensive and repurchasing firms to appear cheaper under EPS-based valuation. In the middle deciles (D4–D7), differences are close to zero, indicating that forecast-type choice matters primarily when expected share-count changes are large. Minor discrepancies can arise because reported EPS forecasts need not equal $\frac{FL_1}{Shares_1}$ exactly due to rounding and differences in analyst coverage across forecast items.[3] Overall, Panel B shows that FPE

[3]For example, for Apple in December 2024, the published EPS forecast is 7.40, whereas dividing FL earnings forecasts by forecasted shares yields 7.37; by contrast, dividing the same FL earnings forecasts by the most recent actual shares yields 7.17. This comparison highlights that expected changes in shares outstanding, rather than rounding or coverage, are the primary driver of differences between EPS- and FL-based FPE ratios.

Table 2.5: Forward P/E ratios: Summary statistics and median differences

Panel A: Summary Statistics for Forward P/E Ratios (2004–2024)

Measure	N	P10	Mean	Median	P90
FPE_{EPS}	117,603	9.3005	20.7272	16.9345	34.3834
FPE_{FL}	117,603	9.2628	25.1498	17.0478	34.0350

Panel B: Median Differences in FPE Ratios by Share Growth Decile (2013–2024)

Decile	$\Delta\text{Shares}_{FY1}$	$\text{Diff}_{\text{EPS-FL}(\%)}$	$\text{FPE}_{\text{1y, EPS}}$	$\text{FPE}_{\text{1y, FL-based}}$	Diff_{FPE}	$\text{Diff}_{\text{FPE}(\%)}$
D1	-6.96%	7.09%	13.6317	14.6866	-0.9759	-6.62%
D2	-4.07%	4.20%	14.9806	15.5907	-0.6229	-4.03%
D3	-2.75%	2.82%	16.8865	17.3389	-0.4742	-2.74%
D4	-1.83%	1.89%	18.1803	18.4619	-0.3357	-1.86%
D5	-1.10%	1.15%	19.2423	19.4306	-0.2216	-1.14%
D6	-0.44%	0.48%	19.9932	20.0277	-0.0935	-0.48%
D7	0.00%	-0.02%	20.5626	20.5144	0.0028	0.02%
D8	0.40%	-0.37%	21.1488	21.1462	0.0714	0.37%
D9	1.31%	-1.23%	21.1606	20.8971	0.2407	1.25%
D10	7.80%	-6.66%	21.4623	19.0406	1.4272	7.10%
N	64,023					

Panel A reports the mean, median, 10th percentile (P10), 90th percentile (P90), and sample size for two variations of forward price-to-earnings ratios, FPE_{EPS} and FPE_{FL}, over 2004–2024. Panel B reports median values by share-growth decile for 2013–2024. Firms are sorted into deciles based on forecasted share growth, $\Delta\text{Shares}_{FY1} = \left(\frac{\text{Forecasted Shares}_{FY1}}{\text{Actual Shares}_{FY0}} - 1\right) \times 100$. $\text{Diff}_{\text{EPS-FL}(\%)}$ denotes $\frac{\text{EPS}_{FY1} - \text{FL}_{FY1}}{\text{FL}_{FY1}} \times 100$. Diff_{FPE} is the difference between FPE ratios based on EPS and FL-based forecasts; $\text{Diff}_{\text{FPE}(\%)}$ expresses this difference as a percentage of the FL-based FPE ratio.

ratios based on EPS and FL forecasts are similar when expected share-count changes are small, but diverge meaningfully for firms with substantial expected repurchases or issuances.

2.4.3 Impact on implied cost of capital

We next examine how the choice between EPS and FL earnings forecasts affects ICC estimates. We compute ICCs using both forecast types and then assess the magnitude, cross-sectional structure, and reliability of the resulting estimates.

Methodology

Because prior research has not established a single ICC model as superior, many studies report a composite ICC constructed as the average of multiple model-based estimates to reduce model-specific noise (e.g., Hou et al. (2012); Li and Mohanram (2014); Gupta (2018); El Ghoul et al. (2018)). Following this approach, we compute ICCs using five widely used models: Gordon (1997; hereafter GG), Easton (2004; MPEG), Claus and

Thomas (2001; CT), Ohlson and Juettner-Nauroth (2005; OJ), and Gebhardt et al. (2001; GLS). The composite ICC is then defined as the simple average of the available model-specific estimates.[4] We exclude ICC estimates below zero and above one and compute the composite ICC from the remaining non-missing model estimates.

To illustrate how forecast type affects ICC estimates, consider the GG model, which depends only on the one-year-ahead earnings input. The EPS-based ICC in this model is

$$\mathrm{ICC}_{GG}^{EPS} = \frac{EPS_1}{P_0}. \tag{2.3}$$

Using the firm-level expressions $P_0 = \frac{MarketCap_0}{Shares_0}$ and $EPS_1 = \frac{FL_1}{Shares_1}$, we can express ICC_{GG}^{EPS} as

$$\mathrm{ICC}_{GG}^{EPS} = \mathrm{ICC}_{GG}^{FL} \cdot \frac{Shares_0}{Shares_1}, \tag{2.4}$$

where

$$\mathrm{ICC}_{GG}^{FL} = \frac{FL_1}{MarketCap_0}.$$

Equation (2.4) shows that for the GG model differences between EPS- and FL-based ICCs arise mechanically from expected changes in shares outstanding. Because the GG ICC is conceptually the inverse of the FPE ratio, the share-count adjustment enters inverted: expected repurchases imply $\frac{Shares_0}{Shares_1} > 1$ and therefore raise ICC_{GG}^{EPS} relative to ICC_{GG}^{FL}, whereas expected issuances imply the opposite. The remaining models incorporate multi-year earnings forecasts and/or explicit growth assumptions, which can amplify the sensitivity of ICC estimates to systematic differences between EPS and FL forecasts over longer horizons.

Empirical ICC estimates

Table 2.6 reports summary statistics for each model-specific ICC and for the composite measure. Across both forecast types, the abnormal earnings growth models $\mathrm{ICC}_{\mathrm{MPEG}}$ and $\mathrm{ICC}_{\mathrm{OJ}}$ have the highest median ICCs (10.00% to 10.59%), followed by the residual income models $\mathrm{ICC}_{\mathrm{GLS}}$ and $\mathrm{ICC}_{\mathrm{CT}}$. The GG model yields the lowest median ICC (around 5.8%), reflecting its no-growth structure. For the composite measure, the median ICC based on EPS is 8.65%, compared with 8.32% using FL forecasts.

Figure 2.2 plots the cross-sectional mean EPS-based ICC estimates for each of the five

[4]Appendix A.2 summarizes the model implementations.

Table 2.6: Summary statistics of ICC estimates

ICC Estimates	Forecast Type	*N*	Mean (%)	Median (%)
Composite	EPS	104,204	9.25%	8.65%
	FL	104,204	9.06%	8.32%
GG	EPS	104,204	6.45%	5.82%
	FL	104,204	6.48%	5.78%
CT	EPS	104,204	8.63%	8.01%
	FL	104,204	8.43%	7.69%
GLS	EPS	104,204	8.45%	8.11%
	FL	104,204	8.43%	8.06%
MPEG	EPS	104,204	11.35%	10.57%
	FL	104,204	10.98%	10.00%
OJ	EPS	104,204	11.36%	10.59%
	FL	104,204	10.99%	10.04%

This table reports descriptive statistics for implied cost of capital (ICC) estimates computed using EPS-based and FL-based earnings forecasts over the period 2004–2024. ICCs are obtained from five standard valuation models: GG, CT, GLS, MPEG, and OJ. A composite ICC measure is calculated as the mean of the five model-specific estimates. For each model and forecast type, the table reports the number of observations (N), along with the mean and median ICC values. EPS-based ICCs use analyst earnings-per-share forecasts, while FL-based ICCs rely on firm-level earnings forecasts scaled by current shares outstanding.

models and for the composite measure. The figure shows substantial level differences across models: the GG model produces consistently lower ICC estimates, reflecting its one-period, no-growth structure and exclusive reliance on one-year-ahead earnings. In contrast, models that incorporate multi-year earnings forecasts and/or explicit growth assumptions (e.g., MPEG, OJ, CT, and GLS) yield higher average ICCs. The composite ICC lies between the individual model estimates and provides a more robust summary measure by averaging across models.

We next turn to the differences between EPS- and FL-based ICC estimates and examine how these wedges vary over time and across the cross-section. ICCs computed from EPS-based forecasts exceed those computed from FL forecasts across all model specifications. To characterize these differences, we compute, for each month, the cross-sectional distribution of the ICC wedge (EPS minus FL).

Figure 2.3 plots the distribution of differences in composite ICC estimates over 2004–2024. The median difference is positive in most months, indicating that EPS-based composite ICCs tend to be systematically higher than FL-based composite ICCs. While the median wedge is typically modest (generally below 0.5 percentage points), the tails are economically meaningful: the upper tail frequently exceeds 1.5 to 2 percentage points (notably around 2009 and again in 2019–2020), while the lower tail reaches negative

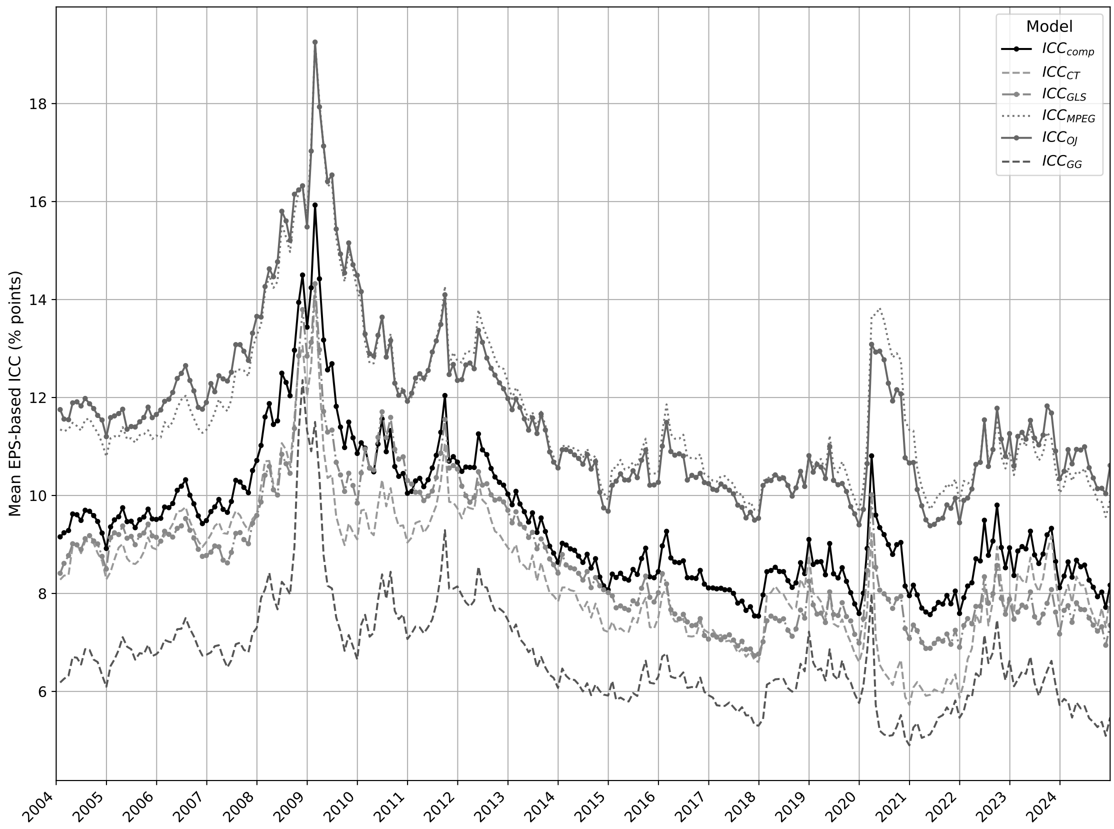


Figure 2.2: Average ICC estimates across models (EPS-based)

This figure plots the cross-sectional mean implied cost of capital (ICC) estimates computed from EPS-based earnings forecasts for five models (GG, MPEG, OJ, CT, and GLS) and for the composite ICC. Each point is the cross-sectional mean at a given forecast date. The sample covers S&P 500 firms over 2004–2024. The composite ICC is the simple average of the available model-specific ICC estimates.

differences of roughly 0.5 to 1 percentage points. The dispersion narrows during episodes of heightened market uncertainty, such as the global financial crisis and the onset of COVID-19, when analysts appear to scale back expectations of share repurchases and differences between EPS- and FL-based forecasts compress.

Table 2.7 complements Figure 2.3 by reporting median ICC differences across deciles sorted on the ICC wedge (EPS minus FL). For the composite ICC, median differences range from -1.40% in the lowest decile to 1.86% in the highest decile. Model-specific results show that MPEG and OJ exhibit the largest tail differences (about -2.26% and -2.29% in decile 1, respectively, and 3.16% and 3.15% in decile 10), CT also shows pronounced tail behavior, while GG and GLS exhibit smaller spreads, consistent with their shorter effective horizons and/or terminal assumptions that attenuate the influence of forecast inputs.

Figure 2.4 further shows that the magnitude of the ICC wedge varies systematically

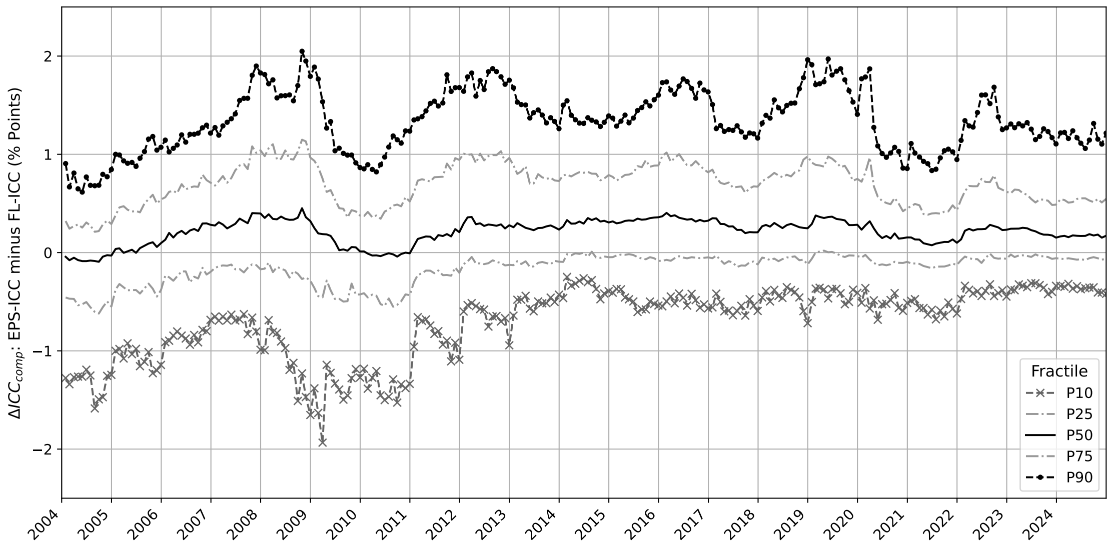


Figure 2.3: Differences in estimated composite ICC

This figure displays the annual cross-sectional 10%, 25%, 50%, 75%, 90% fractiles of the difference (, i.e., ICC_{EPS} - ICC_{FL}) in composite ICC estimates based on the two earnings types, i.e., the composite ICC based on EPS forecasts minus the ICC based on firm-level (FL) forecasts for 2004–2024. The composite ICC is the average of five model estimates: GG (Gordon and Gordon, 1997), GLS (Gebhardt et al., 2001), CT (Claus and Thomas, 2001), MPEG (Easton, 2004), and OJ (Ohlson and Juettner-Nauroth, 2005).

Table 2.7: Median differences between EPS- and FL-based ICC estimates by decile

	ICC Estimates					
Decile	Composite	GG	CT	GLS	MPEG	OJ
1	-1.40%	-0.68%	-1.82%	-0.58%	-2.26%	-2.29%
2	-0.42%	-0.15%	-0.46%	-0.15%	-0.73%	-0.72%
3	-0.16%	-0.05%	-0.15%	-0.05%	-0.28%	-0.27%
4	-0.01%	-0.01%	0.00%	0.00%	-0.03%	-0.03%
5	0.12%	0.02%	0.15%	0.03%	0.19%	0.19%
6	0.27%	0.05%	0.32%	0.07%	0.43%	0.43%
7	0.44%	0.10%	0.54%	0.13%	0.72%	0.71%
8	0.66%	0.17%	0.82%	0.20%	1.09%	1.09%
9	1.02%	0.28%	1.26%	0.31%	1.71%	1.70%
10	1.86%	0.58%	2.30%	0.61%	3.16%	3.15%
N	104,204	104,204	104,204	104,204	104,204	104,204

This table reports median differences between implied cost of capital (ICC) estimates computed using EPS-based forecasts and those computed using FL-based forecasts. Differences are defined as $\text{ICC}_{EPS} - \text{ICC}_{FL}$ and are expressed in percentage points. Firms are sorted into deciles based on $\text{ICC}_{EPS} - \text{ICC}_{FL}$, and medians are reported within each decile. The ICC models considered include Gordon and Gordon (GG), Claus and Thomas (CT), Gebhardt, Lee, and Swaminathan (GLS), Easton's MPEG model, and the Ohlson and Juettner-Nauroth (OJ) model. The Composite ICC measure is the simple average across the five model-specific estimates. The final row reports the number of firm-forecast observations used for each model. The sample period covers 2004–2024.

across models. MPEG, OJ, and CT display the largest spreads, with differences reaching more than three percentage points in some periods, whereas GG and GLS remain comparatively tightly distributed. Across models, median differences are typically positive, indicating that EPS-based ICCs tend to exceed FL-based ICCs.

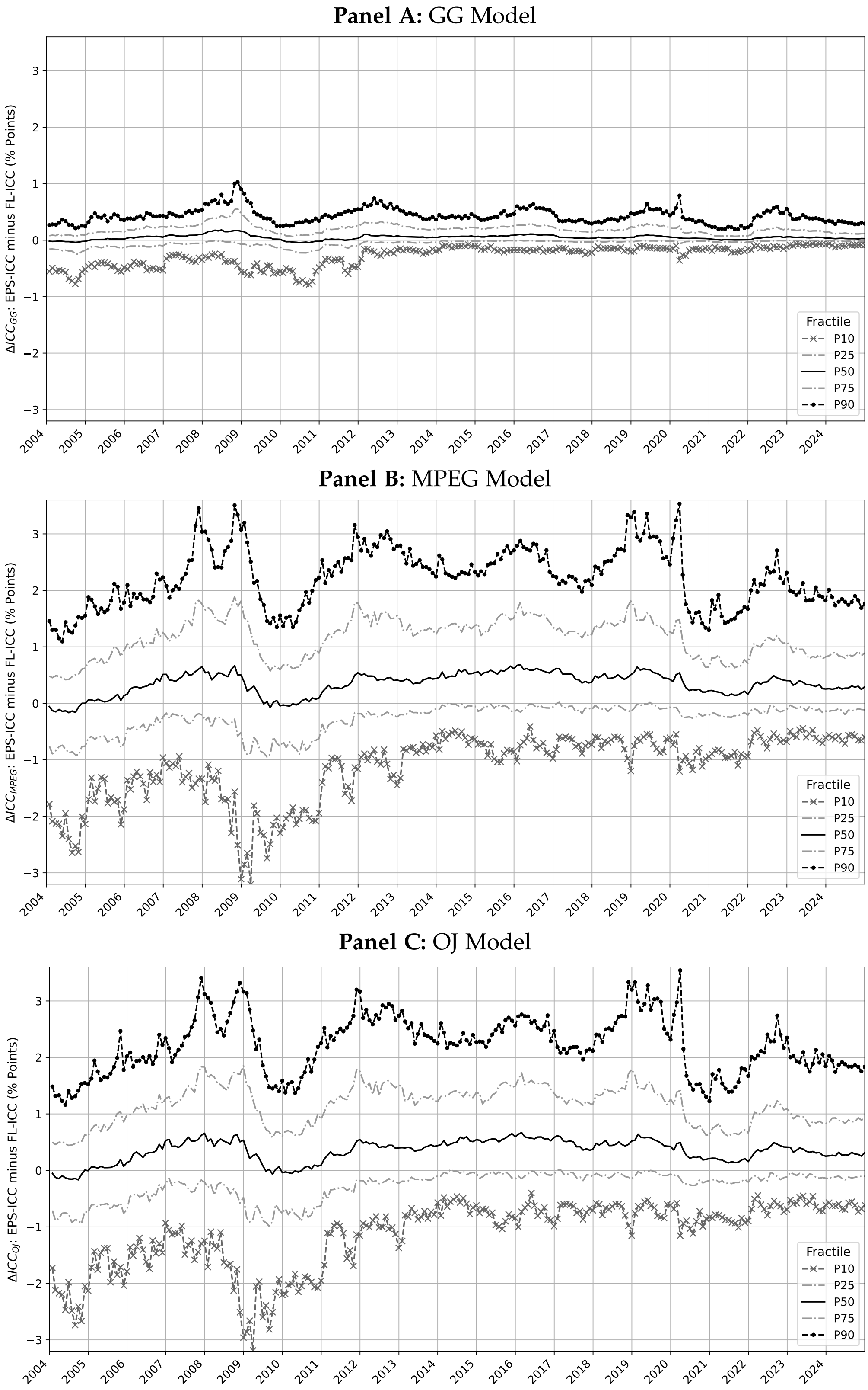
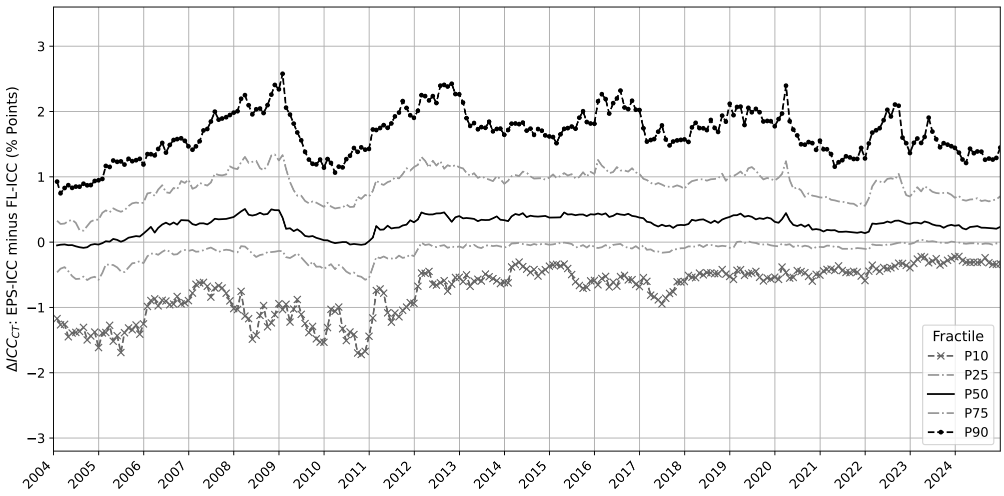


Figure 2.4: Differences in ICC estimates across forecast types per model

Panel D: CT Model



Panel E: GLS Model

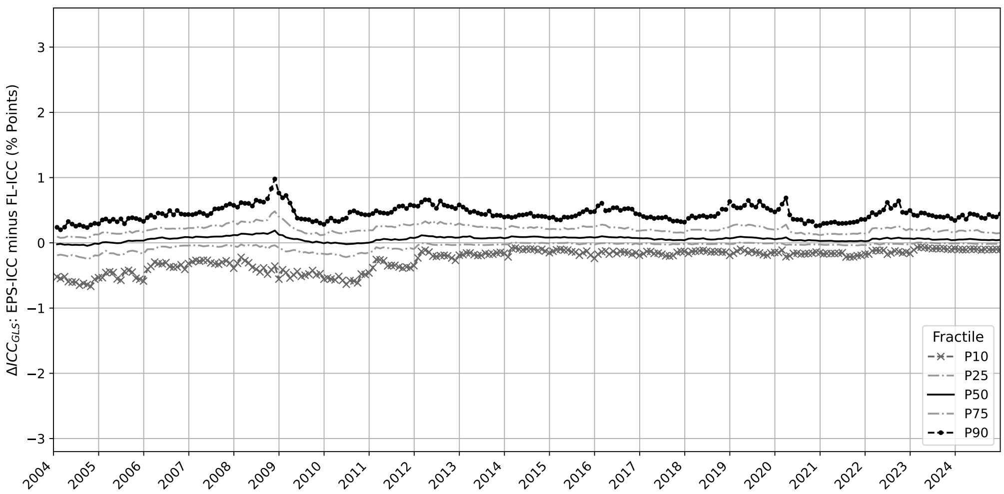


Figure 2.4: Differences in ICC estimates across forecast types per model (continued)
This figure shows the annual cross-sectional 10%, 25%, 50%, 75%, and 90% fractiles of the difference in ICC estimates for each model, computed as EPS-based ICC minus FL-based ICC. ICCs are computed using GG (Gordon and Gordon, 1997) in Panel A, MPEG (Easton, 2004) in Panel B, OJ (Ohlson and Juettner-Nauroth, 2005) in Panel C, CT (Claus and Thomas, 2001) in Panel D, and GLS (Gebhardt et al., 2001) in Panel E. EPS-based ICCs use analyst earnings-per-share forecasts, while FL-based ICCs rely on firm-level earnings forecasts scaled by current shares outstanding. ICCs are calculated for 2004–2024.

Given that ICC differences are largest in the tails, we further sort firms based on the magnitude of the divergence between EPS and FL forecasts and examine how this maps into ICC differences (Table 2.8). Panel A forms deciles based on the difference in one-year-ahead earnings forecasts; firms in the lowest decile have higher FL-based earnings than EPS forecasts (average difference -0.687), whereas firms in the highest decile have higher EPS forecasts (average difference 0.553). Panel B repeats the analysis using five-year-ahead earnings differences, which are considerably larger (from -2.829 in D1 to 3.561 in D10), reflecting increasing divergence at longer horizons. Panels C and D sort on differences in short-term (FY1–FY2) and long-term (FY1–FY5) growth, respectively.

Across these sorts, model sensitivities align with their structures. The GG model is most sensitive to one-year-ahead earnings differences, whereas MPEG and OJ are most sensitive to growth differences, and CT is particularly sensitive to long-horizon earnings and growth differences that affect terminal values. GLS exhibits comparatively small variation. The composite ICC differences remain economically meaningful, ranging from roughly -1.7 to $+1.5$ percentage points across deciles, indicating that aggregating across models does not eliminate sensitivity to forecast type.

Table 2.8: Mean differences in ICCs by decile

Panel / Decile	EPS−FL	Mean ICC Differences by Model (%)					
		Composite	GG	CT	GLS	MPEG	OJ
Panel A: Deciles based on differences in one-year-ahead earnings							
D1	-0.687	-1.33%	-1.25%	-1.62%	-0.75%	-1.53%	-1.52%
D10	0.553	1.38%	0.54%	1.56%	0.49%	2.19%	2.14%
Panel B: Deciles based on differences in five-year-ahead earnings							
D1	-2.829	-1.69%	-0.96%	-2.79%	-0.82%	-1.93%	-1.93%
D10	3.561	1.42%	0.32%	2.11%	0.47%	2.11%	2.07%
Panel C: Deciles based on differences in short-term growth							
D1	-1.089	-1.35%	-0.25%	-1.18%	-0.42%	-2.45%	-2.46%
D10	0.358	1.50%	0.09%	1.32%	0.32%	2.89%	2.89%
Panel D: Deciles based on differences in long-term growth							
D1	-0.098	-1.35%	-0.45%	-2.47%	-0.60%	-1.62%	-1.63%
D10	0.104	1.39%	0.15%	2.37%	0.40%	2.03%	1.99%

This table reports mean differences in ICC estimates based on EPS versus FL forecasts, grouped into deciles. Panels A–D correspond to deciles formed on differences in one-year-ahead earnings, five-year-ahead earnings, short-term growth (FY1 to FY2), and long-term growth (FY1 to FY5), respectively. Only D1 and D10 are shown to highlight extreme effects. ICCs are computed using GG (Gordon and Gordon, 1997), CT (Claus and Thomas, 2001), GLS (Gebhardt et al., 2001), MPEG (Easton, 2004), OJ (Ohlson and Juettner-Nauroth, 2005), and a composite estimate (mean of all five models).

Evaluation of ICCs

Prior research often evaluates expected return proxies by relating them to realized returns. This approach is limited because realized returns are noisy and incorporate unexpected news (Easton and Monahan (2005)). To address this issue, Lee et al. (2021) propose a measurement-error-variance (MEV) framework that evaluates the reliability of expected return proxies without requiring knowledge of the true expected return.

The framework decomposes an expected-return proxy into signal and measurement error and evaluates proxies by the variance of that measurement error. Lee et al. compute average measurement-error variances (SVar) separately for cross-sectional variation (distinguishing expected returns across firms) and time-series variation (tracking within-firm changes over time). Lower SVar implies a more reliable proxy, and SVar can be negative when the covariance between the proxy and realized returns is sufficiently positive. Lee et al. (2021) show that minimizing MEV is necessary and sufficient for selecting the best-performing proxy in a given dimension.

Applying this framework, Table 2.9 reports mean cross-sectional and time-series measurement-error variances (multiplied by 100) for composite ICCs based on EPS and FL forecasts. ICCs constructed with FL forecasts outperform those based on EPS in both dimensions. In the time-series dimension, the FL-based composite ICC has an average SVar of -0.0133, compared to -0.0119 for the EPS-based ICC. In the cross-sectional dimension, the FL-based ICC also exhibits a lower (more negative) SVar (-0.0009 versus 0.0003). Both differences are statistically significant. Overall, these results indicate that FL-based ICCs contain less measurement error and provide a more reliable proxy for expected returns than EPS-based ICCs.

Table 2.9: Measurement Error Variance (Lee, So, and Wang, 2020)

ICC Model	Earnings Type	Time-Series SVar	Cross-Sectional SVar
Composite	EPS	-0.0119***	0.0003
Composite	FL	-0.0133***	-0.0009**
Difference		0.0014**	0.0012*

This table reports scaled measurement-error variances (multiplied by 100) for Composite implied cost of capital (ICC) estimates computed using EPS and FL forecasts. Composite ICCs are the average of five valuation models: GG (Gordon and Gordon, 1997), MPEG (Easton, 2004), CT (Claus and Thomas, 2001), GLS (Gebhardt et al., 2001), and OJ (Ohlson and Juettner-Nauroth, 2005).
The time-series scaled measurement-error variance (Time-Series SVar) is computed as $\text{SVar} = \text{Var}(\widehat{ICC}) - 2\,\text{Cov}(r_{t+1}, \widehat{ICC})$ using one-period-ahead realized returns r_{t+1}, following Lee et al. (2021). The cross-sectional SVar analog uses cross-sectional variance and covariance at each date. SVar can be negative when the covariance term is positive and sufficiently large. Statistical significance is denoted by *, **, and *** for the 10%, 5%, and 1% levels, respectively.

2.4.4 Share repurchases and issuances: firm characteristics and forecast accuracy

Firm characteristics

Identifying the firms most affected by differences between EPS and FL forecasts is important because the resulting wedges can translate into the largest distortions in FPE ratios and ICC estimates. We therefore examine which firm characteristics are associated with expected share repurchases and issuances and whether analysts' expectations are aligned with firms' past share count dynamics.

Table 2.10 relates analysts' expected share changes to historical share count growth. Firms are grouped into deciles based on past growth in shares outstanding, where the lowest decile contains firms with the largest historical share reductions (net repurchasers) and the highest decile contains firms with the largest historical share increases (net issuers). Forecasted share changes broadly follow historical patterns, but expectations are more muted. For example, at the one-year horizon, firms with the largest historical buybacks experienced an average share decline of 7.33%, whereas analysts forecast a decline of 4.79%; firms with the largest historical issuances experienced an average share increase of 9.61%, whereas analysts forecast an increase of only 2.82%. The asymmetry is even stronger over a three-year window: historical issuers increased shares by 10.74%, but analysts forecast essentially flat share counts (0.42%), while historical repurchasers are still expected to reduce shares materially (3.45%). Overall, analysts anticipate substantially less persistence in issuances than in repurchases, consistent with repurchase programs being more predictable and often pre-announced, whereas equity issuances are frequently opportunistic and harder to time ex ante (Asquith and Mullins Jr (1986)).

We next examine how firm characteristics vary across firms with different equity payout and financing policies. Table 2.11 reports average firm characteristics across deciles sorted by actual (Panel A) and forecasted (Panel B) changes in shares outstanding. Decile 1 contains net repurchasers (largest share reductions) and decile 10 contains net issuers (largest share increases). The table reports sales growth, stock returns, size, leverage, profitability, free cash flow, stock-based compensation, and lagged share growth. Characteristics are measured three months after fiscal year-end; in Panel A they are additionally lagged by one year to avoid look-ahead bias. The Spread_{10-1} row reports differences between issuers and repurchasers, while $\text{Spread}_{10/1-5/6}$ compares the tails to

Table 2.10: Historical and forecasted growth in number of shares outstanding

Decile	1-Year		3-Year	
	Hist (%)	Frc (%)	Hist (%)	Frc (%)
1	-7.33	-4.79	-6.25	-3.45
2	-4.32	-3.42	-3.91	-2.58
3	-2.94	-2.44	-2.69	-1.95
4	-2.02	-1.77	-1.85	-1.57
5	-1.24	-1.15	-1.12	-1.07
6	-0.55	-0.48	-0.44	-0.50
7	0.08	0.02	0.18	-0.05
8	0.56	0.26	0.90	0.26
9	1.82	0.91	2.98	0.66
10	9.61	2.82	10.74	0.42

This table presents 1-year and 3-year historical and forecasted growth rates in the number of shares outstanding by decile. Firms are sorted into deciles based on their historical change in shares outstanding for each respective period. "Hist" refers to the realized historical growth rate, while "Frc" denotes the forecasted rate derived from analysts' shares-outstanding forecasts. All values are expressed in percent.

the center to capture potential non-linearities.

The strongest monotonic pattern concerns growth and profitability. In both panels, sales growth is substantially higher for issuers than for repurchasers: in Panel A it increases from 4.4% in decile 1 to 8.8% in decile 10 (spread 4.4 percentage points), and Panel B shows a similar spread of 4 percentage points. In contrast, profitability and internal liquidity decline sharply from repurchasers to issuers. In Panel A, RoA falls from 11.8% (decile 1) to 5.2% (decile 10), and free cash flow declines from 10.9% to 2.9%; Panel B exhibits comparable spreads (RoA: 10.9% to 4.8%; FCF: 10.2% to 3.5%). These patterns suggest that issuers tend to be fast-growing but less profitable and more reliant on external financing, whereas repurchasers tend to be mature, cash-generative, and highly profitable firms.

Several other characteristics exhibit non-linear patterns. The $\text{Spread}_{10/1-5/6}$ indicates U-shaped relationships for returns, size, leverage, and stock-based compensation: firms in the tails (large repurchasers and large issuers) differ from firms with relatively stable share counts. In addition, issuers are more leveraged than repurchasers (Spread_{10-1} of 3.7 p.p. in Panel A and 3.1 p.p. in Panel B), consistent with the view that equity issuance can relieve financing constraints. Overall, analysts' expectations appear closely aligned with observable firm fundamentals, helping to identify settings in which EPS-based performance measures and valuation inputs are most likely to be affected by embedded share-count expectations.

Table 2.11: Firm characteristics across deciles sorted by changes in shares outstanding

Decile	ΔShares	Growth	RET	Size	LEV	RoA	FCF	Stock Comp	$\Delta\text{Shares}_{t-1}$
Panel A: Actual Shares									
1	-7.1%	4.4%	9.4%	9.89	28.4%	11.8%	10.9%	0.39%	-4.8%
2	-4.0%	5.0%	7.9%	10.05	25.3%	11.7%	8.4%	0.35%	-3.1%
3	-2.6%	4.8%	9.1%	9.89	23.9%	12.1%	7.5%	0.33%	-2.2%
4	-1.8%	6.0%	11.4%	10.21	24.9%	12.6%	6.8%	0.34%	-1.6%
5	-1.1%	5.7%	11.6%	10.05	24.4%	12.3%	6.1%	0.32%	-1.2%
6	-0.5%	6.3%	12.4%	9.90	24.8%	11.0%	5.2%	0.31%	-0.5%
7	0.1%	5.1%	8.8%	9.82	29.5%	8.3%	4.9%	0.28%	0.1%
8	0.6%	6.2%	10.8%	9.82	28.3%	7.2%	4.5%	0.35%	0.5%
9	2.0%	8.3%	12.8%	9.93	30.2%	5.9%	3.2%	0.35%	1.4%
10	10.1%	8.8%	7.5%	9.78	32.0%	5.2%	2.9%	0.38%	3.5%
Spread_{10-1}	0.1718***	0.0440***	-0.0194	-0.1142	0.0366**	-0.0664***	-0.0794***	-0.0001	0.0830***
$\text{Spread}_{10/1-5/6}$	0.0226***	0.0058	-0.0355*	-0.1390*	0.0562***	-0.0319***	0.0125***	0.0007***	0.0018
Panel B: Forecasted Shares									
1	-7.4%	3.8%	7.8%	9.94	28.9%	10.9%	10.2%	0.40%	-4.2%
2	-4.0%	4.0%	9.8%	10.04	28.4%	11.6%	8.1%	0.35%	-3.3%
3	-2.6%	4.5%	9.4%	10.12	25.9%	11.6%	7.5%	0.32%	-2.2%
4	-1.8%	4.6%	10.5%	10.17	25.5%	11.8%	6.7%	0.33%	-1.7%
5	-1.2%	5.3%	12.2%	10.09	26.8%	11.8%	6.1%	0.31%	-1.1%
6	-0.6%	5.6%	14.3%	10.07	26.0%	11.4%	5.3%	0.29%	-0.6%
7	0.0%	5.6%	11.2%	9.99	28.4%	9.1%	4.7%	0.27%	0.0%
8	0.6%	5.6%	10.5%	9.94	29.1%	7.4%	4.5%	0.33%	0.3%
9	2.0%	6.8%	11.6%	10.04	29.3%	5.8%	3.4%	0.36%	1.0%
10	9.6%	7.8%	5.8%	9.84	32.1%	4.8%	3.5%	0.40%	2.3%
Spread_{10-1}	0.1698***	0.0401***	-0.0199	-0.0992	0.0313**	-0.0612***	-0.0670***	-0.0000	0.0649***
$\text{Spread}_{10/1-5/6}$	0.0196***	0.0035	-0.0642***	-0.1904***	0.0412***	-0.0372***	0.0113**	0.0010***	-0.0008

This table reports average firm characteristics across deciles sorted by actual and expected changes in shares outstanding. Panel A reports values based on actual changes in shares, $(\text{NoSh\_Dil\_L0}/\text{NoSh\_Dil\_L1}) - 1$, while Panel B reflects forecasted changes, $(\text{Frc\_Nosh\_F1}/\text{Act\_Nosh\_L0}) - 1$. Size is the natural logarithm of market capitalization. RoA is EBIT over total assets. Growth is annual sales growth. RET is annual stock return. Stock Comp is stock-based compensation scaled by market equity. LEV is long-term debt over equity. FCF is free cash flow scaled by market equity. $\Delta\text{Shares}_{t-1}$ denotes the percentage change in shares outstanding between years $t-1$ and t. All firm-characteristic variables in Panel A are lagged by one year. The characteristics are calculated three months after each fiscal year-end from 2014–2024, during which forecasted share data are available. Spread_{10-1} captures the difference between the top and bottom decile. $\text{Spread}_{10/1-5/6}$ compares extreme deciles to the middle to account for potential U-shaped relationships. ***, **, and * denote 1%, 5%, and 10% significance levels.

Table 2.12 provides complementary evidence using rank correlations. In Panel A (actual data), ΔShares is positively correlated with sales growth (0.08) and returns (0.06), and strongly correlated with lagged share changes (0.57), indicating persistence in issuance and repurchase behavior. At the same time, ΔShares is negatively correlated with profitability (RoA = -0.37) and free cash flow (FCF = -0.32), implying that firms with stronger operating performance and internal liquidity are more likely to repurchase shares. Panel B (forecasted data) mirrors these patterns: expected share growth is positively related to sales growth (0.15) and strongly related to lagged share changes

(0.67), but negatively related to RoA (-0.41) and FCF (-0.38). One notable difference is that stock-based compensation is negatively correlated with actual share changes (-0.10) but shows no significant relation with expected share changes, suggesting that analysts underweight dilution related to equity compensation when forming share-count forecasts.

Table 2.12: Spearman rank correlation matrices: actual and forecasted data

Panel A: Actual Data

	ΔShares	Growth	RET	Size	LEV	RoA	FCF	StockComp	$\Delta\text{Shares}_{t-1}$
ΔShares									
Growth	0.08***								
RET	0.06***	0.10***							
Size	-0.08***	0.07***	0.13***						
LEV	0.08***	-0.15***	-0.01	0.04***					
RoA	-0.37***	0.18***	0.07***	0.10***	-0.16***				
FCF	-0.32***	-0.08***	-0.12***	-0.07***	0.06***	0.20***			
StockComp	-0.10***	0.04**	-0.07***	-0.17***	-0.11***	-0.10***	0.06***		
$\Delta\text{Shares}_{t-1}$	0.57***	0.32***	0.10***	-0.04**	-0.00	-0.32***	-0.29***	-0.08***	

Panel B: Forecast Data

	$\Delta\text{Shares}_{t+1}$	Growth	RET	Size	LEV	RoA	FCF	StockComp	ΔShares
$\Delta\text{Shares}_{t+1}$									
Growth	0.15***								
RET	0.03	0.19***							
Size	-0.06**	0.11***	0.15***						
LEV	0.03	-0.08***	-0.06***	-0.04*					
RoA	-0.41***	0.12***	0.17***	0.12***	-0.05**				
FCF	-0.38***	-0.06***	-0.09***	-0.08***	0.05**	0.20***			
StockComp	0.00	0.04*	-0.13***	-0.27***	-0.11***	-0.16***	0.12***		
ΔShares	0.67***	0.33***	0.07***	0.00	0.02	-0.36***	-0.32***	-0.04*	

This table reports Spearman rank correlations among key firm-level variables. Panel A uses realized data, where ΔShares denotes the percentage change in shares outstanding between years $t-1$ and t. Panel B uses forecasted data, where $\Delta\text{Shares}_{t+1}$ reflects expected changes in shares outstanding from year t to $t+1$. Growth is annual sales growth; RET is annual stock return; Size is the natural logarithm of market capitalization; LEV is long-term debt over equity; RoA is EBIT over total assets; FCF is free cash flow scaled by market equity; StockComp is stock-based compensation scaled by market equity. In Panel A, firm characteristics are lagged by one year. ***, **, and * denote significance at the 1%, 5%, and 10% levels.

To assess whether these relations persist in a multivariate setting, Table 2.13 reports panel regressions of forecasted and realized share changes on firm characteristics. The estimates reinforce the univariate evidence: sales growth is positively associated with issuance, while profitability (RoA) and free cash flow are negatively associated with share growth, consistent with repurchases among cash-rich firms. Leverage is positively associated with realized share growth, indicating that more indebted firms are more likely to issue equity. Stock-based compensation is negatively related to realized share growth (but not to expected share growth), consistent with firms repurchasing shares to offset dilution from option exercises and equity compensation (Jolls (1998)). Including lagged share changes improves explanatory power, consistent with persistence in equity financing and payout policies.

Taken together, the evidence indicates that share issuers tend to be fast-growing, more leveraged, and less profitable firms with weaker internal cash generation, whereas repurchasers tend to be highly profitable, cash-rich firms with relatively low sales growth. This profile is consistent with equity issuance to finance investment and relieve constraints (see Fama and French (2005)), and with repurchases as a dominant payout policy among mature firms. Importantly, because repurchases mechanically increase EPS by reducing shares outstanding, these patterns also highlight settings in which EPS-based performance measures can diverge most from firm-level earnings. Consistent with an EPS-management interpretation, repurchasing firms combine low growth with high profitability and strong cash flows, suggesting that repurchases may be used not only to return capital but also to support per-share performance metrics (see Hribar et al. (2006); Almeida et al. (2016)). At the same time, the weaker role of leverage and stock-based compensation in explaining forecasted (relative to realized) share changes suggests that analysts underweight financing structure and dilution-related incentives when forming expectations about future share activity.

Table 2.13: Regression results of forecasted and actual change in shares outstanding

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10
Panel A: Actual Share Changes $(act\_nosh\_diff = \frac{shares(t)}{shares(t-1)} - 1)$										
Growth	0.022***								0.044***	0.004
RET		0.008***							0.005**	0.002
Size			0.000						0.002***	0.001***
LEV				0.024***					0.031***	0.026***
RoA					-0.117***				-0.195***	-0.125***
FCF						-0.125***			-0.103***	-0.085***
StockComp							-0.533**		-0.867***	-0.469**
ΔShares$_{t-1}$								0.338***		0.287***
R^2 (Within)	0.001	0.002	-0.000	0.003	0.024	0.038	0.003	0.008	0.062	0.059
R^2 (Between)	0.042	0.016	0.006	0.017	-0.023	-0.028	-0.021	0.449	0.208	0.467
R^2 (Overall)	0.008	0.002	-0.001	-0.003	0.013	0.024	-0.001	0.129	0.107	0.175
Observations	5,898	5,898	5,898	5,898	5,898	5,898	5,898	5,898	5,898	5,898
Panel B: Analyst Forecasts $(Frc[shares(t+1)]/shares(t) - 1)$										
Growth	0.007								0.012	0.004
RET		0.002							0.003	0.001
Size			-0.000						0.002***	0.001***
LEV				0.001					0.003	-0.001
RoA					-0.077***				-0.126***	-0.080***
FCF						-0.102***			-0.089***	-0.095***
StockComp							-0.110		-0.314	-0.243
ΔShares$_{t-1}$								0.200***		0.174***
R^2 (Within)	-0.000	0.000	-0.000	0.000	0.012	0.022	0.000	-0.024	0.028	0.005
R^2 (Between)	0.013	-0.002	0.000	-0.000	0.023	0.035	-0.001	0.316	0.117	0.361
R^2 (Overall)	0.002	-0.000	0.000	-0.000	0.013	0.027	0.000	0.052	0.050	0.085
Observations	3,300	3,300	3,300	3,300	3,300	3,300	3,300	3,300	3,300	3,300

This table presents panel regression results with random effects and clustered standard errors. The dependent variable is the change in shares outstanding. Panel A reports actual changes $(\text{NoSh\_Dil\_L0}/\text{NoSh\_Dil\_L1}) - 1$, and Panel B reports forecasted changes $(\text{Frc\_Nosh\_F1}/\text{Act\_Nosh\_L0}) - 1$. In Panel A, firm characteristics are lagged by one year. Growth is annual sales growth; RET is annual stock return; Size is the natural logarithm of market capitalization; LEV is long-term debt over equity; RoA is EBIT over total assets; FCF is free cash flow scaled by market equity; StockComp is stock-based compensation scaled by market equity. Panel A covers 2004–2024; Panel B covers 2013–2024. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Forecast accuracy

Table 2.14 evaluates the accuracy and efficiency of analysts' forecasts of share count changes and examines whether forecast errors are systematically related to firm characteristics.

Panel A reports Mincer–Zarnowitz regressions of realized share changes on forecasted share changes. Slope coefficients are close to one across horizons, indicating that forecasts track realized share changes reasonably well on average. At the same time, significant intercepts and rejections of forecast efficiency point to mild systematic biases: analysts slightly underpredict near-term share changes and overpredict longer-horizon activity. The declining R^2 values across horizons further indicate that forecast accuracy deteriorates at longer horizons.

Panel B confirms these patterns in specifications with firm and time fixed effects. Forecasts reliably explain realized share changes at the one- and two-year horizons, but their explanatory power weakens at longer horizons, consistent with analysts becoming more conservative and less able to anticipate longer-term equity financing and payout decisions.

Panel C relates forecast errors to firm characteristics. Most coefficients are insignificant, suggesting that analysts incorporate much of the observable information relevant for predicting issuance and repurchases. However, two variables appear to be systematically underweighted. Leverage is positively associated with forecast errors across horizons, implying that highly levered firms issue more shares (or repurchase fewer) than analysts anticipate. Stock-based compensation is significantly negative at the one-year horizon, indicating that firms with high equity-based pay tend to repurchase more than forecasted, consistent with repurchases used to offset dilution from option exercises. Profitability is negatively associated with forecast errors at shorter horizons, suggesting that profitable firms repurchase more than anticipated.

Overall, these results indicate that analysts' share count forecasts are broadly accurate but not fully efficient. While forecasts reflect key fundamentals such as growth and profitability, analysts underreact to leverage and stock-based compensation, two determinants of realized share changes that plausibly reflect financing constraints and dilution-management incentives.

Finally, we examine whether incorporating expected share count changes into EPS forecasts improves earnings forecast accuracy. Table 2.15 reports descriptive statistics

Table 2.14: Forecast vs. actual share changes: regression of shares and forecast errors

Statistic	L=1	L=2	L=3
Panel A: MZ Regression of Actual Shares on Forecasted Shares			
Intercept	-0.0010**	0.0034***	0.0060***
Slope	0.9498***	1.0784***	1.0430***
F-stat ($b_0 = 0$ & $b_1 = 1$)	17.01***	20.33***	32.82***
R^2	0.7473	0.5430	0.4139
Panel B: Panel Regression of Actual Shares on Forecasted Shares			
Slope	0.9520	0.9521	0.7719
Wald test	0.2002	0.1463	0.0014
R^2 (within)	0.7311	0.4153	0.2331
Panel C: Regression of Forecast Error (FE) on Characteristics			
Growth	-0.0023	0.0106	0.2017
RET	0.0001	-0.0058	-0.0380**
Size	-0.0010	-0.0188	-0.0264
LEV	0.0135***	0.0296**	0.0699*
FCF	-0.0077	0.0245	0.0159
RoA	-0.0259*	-0.0909**	-0.1027
StockComp	-0.3106*	-2.3835	-2.8347
ΔShares	0.0096*	0.2110	0.2947
Adj. R^2	0.0110	0.0364	0.0549

This table presents regression results examining the relationship between forecasted and actual changes in firms' shares outstanding over forecast horizons of one (L=1), two (L=2), and three years (L=3) between 2013 and 2024. Panel A shows Mincer–Zarnowitz (MZ) regressions of actual share changes on forecasted share changes. Panel B reports panel regressions of actual share changes on forecasted share changes with firm and time fixed effects; the Wald test evaluates whether the slope differs from one. Panel C reports pooled regressions of forecast errors (FE) on firm characteristics: Growth (sales growth), RET (stock return), Size (log market capitalization), LEV (leverage), FCF (free cash flow scaled by market equity), RoA (operating profitability), StockComp (stock-based compensation scaled by market equity), and ΔShares (actual change in shares outstanding). Significance: ***, **, and * denote 1%, 5%, and 10% levels, respectively.

for forecast errors across horizons under alternative share-scaling conventions. Panel A reports share-scaled forecast errors for share count forecasts, while Panels B and C report price-scaled forecast errors (PFE) and absolute price-scaled forecast errors (PAFE) for earnings forecasts under different share bases. Across horizons, mean PFEs based on reported EPS show slightly higher bias than some alternatives, but EPS-based forecasts deliver the smallest absolute errors: EPS-based PAFEs have the lowest mean and median values relative to approaches that scale firm-level forecasts using lagged shares or realized shares. These results indicate that reported EPS forecasts best capture analysts' intended per-share expectations and that alternative reconstructions introduce additional noise.

Table 2.16 confirms these patterns using paired tests. EPS-based PAFEs are statistically smaller than those based on lagged-share scaling, and EPS-based absolute errors are also slightly but significantly smaller than those based on realized-share ("cheat") scaling

Table 2.15: Summary statistics of forecast errors by type and horizon

Error Type	Share Basis	FY1		FY2		FY3	
		Mean	Median	Mean	Median	Mean	Median
Panel A: Share-scaled Forecast Errors (SFE/SAFE)							
SFE	Diluted	-0.0019	-0.0005	0.0067	-0.0003	0.0205	0.0031
SAFE	Diluted	0.0143	0.0055	0.0375	0.0153	0.0633	0.0270
Panel B: Price-scaled Forecast Errors (PFE)							
PFE	EPS	-0.0004	0.0008	-0.0026	0.0000	-0.0034	-0.0004
	Cheat	-0.0004	0.0007	-0.0026	0.0000	-0.0028	-0.0003
	Lagged	0.0004	0.0013	-0.0011	0.0008	-0.0007	0.0010
	Forecast	-0.0003	0.0008	-0.0026	0.0001	-0.0036	-0.0003
Panel C: Price-scaled Absolute Forecast Errors (PAFE)							
PAFE	EPS	0.0106	0.0035	0.0187	0.0070	0.0238	0.0102
	Cheat	0.0108	0.0035	0.0191	0.0075	0.0243	0.0106
	Lagged	0.0112	0.0041	0.0194	0.0079	0.0249	0.0112
	Forecast	0.0106	0.0035	0.0188	0.0071	0.0242	0.0100
N		3,673		3,163		2,485	

This table reports mean and median forecast errors by error type, share basis, and forecast horizon (FYi) for 2013–2024. **Share-scaled errors (SFE/SAFE)** measure the difference between actual and forecasted shares outstanding scaled by lagged diluted shares, capturing signed and absolute deviations, respectively. **Price-scaled forecast errors (PFE)** measure the difference between actual and forecasted EPS scaled by stock price; **price-scaled absolute forecast errors (PAFE)** are the absolute values. Share adjustments convert firm-level forecasts into per-share terms: **EPS** uses reported EPS, **Cheat** divides firm-level forecasts by actual diluted shares in FYi, **Lagged** uses lagged shares, and **Forecast** uses forecasted shares outstanding.

across horizons. At the same time, EPS-based forecasts exhibit greater bias than FL forecasts scaled by lagged shares, consistent with EPS embedding share-count expectations that affect the level and growth interpretation. Overall, the evidence is consistent with Kaplan et al. (2024): incorporating expected share-count changes into EPS forecasts improves forecast accuracy on an absolute-error basis, even though it can distort the interpretation of earnings growth and valuation inputs when EPS is used as a proxy for FL earnings.

2.5 Conclusion

This paper examines an underexplored distinction in financial analysis: the difference between earnings per share and firm-level earnings forecasts, and its implications for commonly used valuation metrics. While analysts' EPS forecasts are widely used in valuation and asset pricing, we document that embedded assumptions about future share counts—reflecting expected repurchases or issuances—can mechanically distort

Table 2.16: Mean differences in forecast errors

Error Type	Horizon	Forecast Pair	Mean Difference
Panel A: Price-scaled Forecast Errors (PFE)			
PFE	FY1	EPS − Lagged	-0.0008***
		EPS − Cheat	0.0000
	FY2	EPS − Lagged	-0.0015***
		EPS − Cheat	-0.0001
	FY3	EPS − Lagged	-0.0027***
		EPS − Cheat	-0.0007***
Panel B: Price-scaled Absolute Forecast Errors (PAFE)			
PAFE	FY1	EPS − Lagged	-0.0005***
		EPS − Cheat	-0.0001*
	FY2	EPS − Lagged	-0.0007***
		EPS − Cheat	-0.0004***
	FY3	EPS − Lagged	-0.0011***
		EPS − Cheat	-0.0005**

This table reports mean differences in forecast errors between EPS-based forecasts and alternative share-basis constructions for horizons FY1–FY3. "Lagged" scales firm-level forecasts by lagged diluted shares; "Cheat" scales firm-level forecasts by realized diluted shares in FYi. Mean differences are computed as (EPS minus alternative) and reported separately for price-scaled forecast errors (PFE) and price-scaled absolute forecast errors (PAFE). ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

EPS-implied growth, forward price-earnings ratios, and implied cost of capital estimates. These assumptions can improve EPS forecast accuracy, but they also blur the interpretation of EPS-based signals by mixing operating earnings expectations with expected payout and financing policy.

Importantly, these distortions are not random in the cross-section. They align with systematic differences in firm characteristics, including growth opportunities, profitability, and stock-based compensation. While concurrent work shows that explicit share-count forecasts can improve EPS forecast accuracy (Bradshaw et al. (2025); Kaplan et al. (2024)), our evidence implies that greater EPS accuracy need not translate into a cleaner signal of expected operating performance because EPS forecasts mechanically incorporate expected share-count changes.

Our findings suggest several directions for future research. First, analysts' forecasts of share-count changes may carry information beyond what is embedded in earnings forecasts, reflecting expectations about payout policy, financing needs, and managerial incentives. Understanding what drives these forecasts and what they reveal about firm behavior would be a natural extension of this work. A second direction is to study whether the recent shift from cash dividends toward repurchases is linked to the composition of stock-based compensation packages (e.g., options versus restricted stock and performance-vested awards), and whether these incentives strengthen firms' emphasis on managing per-share performance measures.

Chapter 3

From Street to GAAP: Forecast Definitions in Implied Cost of Capital Estimation†

3.1 Introduction

A large share of capital markets research relies on analyst earnings forecasts as a forward-looking summary of market expectations. In most empirical settings, however, the analyst forecast is implicitly treated as a Street (non-GAAP) forecast. This convention is not accidental. A substantial literature argues that Street earnings are more value relevant than GAAP earnings in short-window pricing tests and appear closer to what investors price when interpreting performance (e.g., Brown and Sivakumar (2003); Kolev et al. (2008); Bradshaw and Sloan (2002)). As a result, implied cost of capital (ICC) studies almost uniformly construct ICCs using Street forecasts, and the earnings-definition choice is rarely discussed as a modeling decision.[1]

The historical evolution of the data reinforces this convention. I/B/E/S began providing explicit bottom-line GAAP earnings forecasts only in 2003. Earlier ICC implementations therefore relied on the more widely available Street forecast inputs, and the empirical literature developed around the idea that non-GAAP measures may better capture persistent core performance. Consistent with this view, short-window announce-

† This paper is based on Zährl (2026). I thank Dieter Hess for helpful comments and suggestions.

[1] To the best of my knowledge, no study using analyst forecast data constructs ICCs from GAAP earnings forecasts. The only exception in the ICC literature are cross-sectional earnings models (e.g.,Hou et al. (2012)) which use Compustat-based GAAP earnings by construction.

ment evidence documents that pro forma earnings can be more informative to investors than GAAP earnings in explaining abnormal returns around earnings releases (Bhattacharya et al. (2003)), and prior work shows that market reactions and earnings response coefficients often appear stronger when earnings are aligned with Street definitions (Bradshaw and Sloan (2002)). Together, limited GAAP forecast availability and early pricing evidence contributed to the widespread practice of treating Street forecasts as the natural earnings input in forward-looking applications such as ICC estimation.

More recent work, however, emphasizes that GAAP-versus-Street comparisons are not straightforward and that conventional designs can obscure the information content of GAAP expectations. First, much of the early evidence is affected by a measurement-error and benchmarking problem: because explicit GAAP expectations were historically unavailable, studies often approximated GAAP forecast errors using constructs that embed substantial noise and definitional mismatches between realized and expected earnings. Bradshaw et al. (2018) exploit newly available analyst GAAP forecasts to show that traditionally identified GAAP forecast errors contain economically large measurement error, and that mitigating this error changes inference in prominent settings. Related research similarly argues that apparent differences in GAAP versus Street earnings response coefficients can be driven by errors-in-variables arising from mismatched definitions of realized and expected earnings (Cohen et al. (2007)).

Second, even when expectations are measured correctly, GAAP and Street earnings are mechanically linked: Street earnings equal GAAP earnings net of firm-specific exclusions. The Street–GAAP forecast difference therefore reflects an implicit expectation about excluded items. Empirical specifications that place GAAP and non-GAAP measures side-by-side without disciplining the decomposition can blur the distinction between the accounting-consistent component and the incremental adjustment component. Christensen et al. (2025) emphasizes that isolating the adjustment component is central for identification and for interpreting coefficients in the presence of this mechanical linkage.

This issue naturally carries over to ICC estimation. ICCs infer an expected return from a valuation equation that uses earnings forecasts at multiple horizons, so the relevant inputs are the joint configuration of forecast levels and implied growth and how a given ICC model weights short-run versus long-run components. If Street and GAAP forecasts differ along these dimensions, ICCs constructed from Street forecasts and those constructed from GAAP forecasts can differ even when prices and implementation choices are held fixed, and the sign and magnitude of the difference can vary across valuation models.

Despite the central role of analyst forecasts in ICC research, there is little evidence that directly evaluates how the choice of earnings definition affects ICC construction and expected-return inference, given that the mechanical linkage between GAAP and Street measures is respected. The analysis therefore revisits the Street-versus-GAAP question through the lens of ICCs. Using matched Street and GAAP analyst earnings forecasts for S&P 500 firms from 2004 to 2024, ICC estimates are constructed under each definition for five widely used ICC models and for a composite measure that averages across models. Prices, book values, and implementation choices are held constant across definitions, so differences in implied discount rates arise exclusively from differences in forecast inputs. Definition effects are summarized by the Street–GAAP ICC differential, i.e., the difference between ICCs constructed from Street versus GAAP forecast inputs.

The descriptive evidence highlights the economic mechanism. Street forecasts tend to be higher in levels, while GAAP forecasts frequently imply steeper growth across horizons. Since valuation models weight these components differently, definition effects are model dependent. Growth-sensitive models translate growth differences into particularly large ICC differentials, with median magnitudes in excess of three percentage points in the left tail of the growth-difference distribution. Even when Street- and GAAP-based ICCs are highly correlated, definition choice reshuffles a non-trivial fraction of firms across the tails of the ICC distribution. For example, top-decile turnover reaches 27.7% for MPEG and remains economically meaningful even for the composite ICC (21.6%). Differences also persist over time in the cross-sectional tails of the composite Street–GAAP ICC differential distribution.

The analysis has two components. First, measurement-error diagnostics following Lee et al. (2021) assess the reliability of Street- and GAAP-based ICC estimates. In this step, both proxies display broadly similar reliability in the aggregate, suggesting that each contains economically meaningful return-relevant signal and that differences in subsequent return results are unlikely to be driven solely by first-order noise differences.

Second, the paper examines whether GAAP- and Street-based ICCs differ in their association with subsequent realized returns. Because Street forecasts mechanically combine a GAAP-consistent expectation with an incremental adjustment component, similar aggregate reliability does not imply that Street- and GAAP-based ICCs reflect the same underlying information. Guided by this decomposition, the analysis relates GAAP-based ICCs and the Street–GAAP ICC differential to subsequent realized returns and studies where these relations are strongest. The evidence is concentrated at short horizons and varies meaningfully across valuation models. At the 12-month horizon, GAAP-based ICCs are positively associated with subsequent returns in most implementations, with particularly strong effects for the model of Claus and Thomas (2001) and for the composite measure. The Street–GAAP ICC differential is also positively associated with subsequent returns in several models, but the pattern is less uniform and weak in the model of Gebhardt et al. (2001). These relations attenuate at longer horizons and are generally small by 36 months. Cross-sectional heterogeneity further indicates stronger GAAP-based ICC effects when forecast-definition divergence is high and in settings linked to greater informational frictions, most notably smaller firms and intangible-intensive industries, suggesting incremental return-relevant information in GAAP-based expectations even in samples where Street-based ICCs remain the default.

Most ICC research treats the analyst-forecast definition as a secondary implementation choice, typically defaulting to Street forecasts, and focuses instead on differences across valuation models or the use of composites. The paper's contribution is to treat forecast definition as a first-order element of ICC measurement and expected-return inference. By varying only the earnings forecast definition while holding prices and non-forecast inputs fixed, the analysis provides a clean mapping from definitional differences in forecast levels and implied growth into implied discount-rate estimates across widely used ICC implementations. This framework clarifies why model choice and composite construction need not eliminate sensitivity to forecast inputs and why forecast definition can matter for empirical designs that rely on ICC rankings and tail assignments. More broadly, the evidence supports treating GAAP-based ICCs as viable expected-return proxies and suggests that incorporating GAAP expectations is particularly informative when Street and GAAP forecasts diverge and in settings associated with greater informational frictions.

The remainder of the paper proceeds as follows. Section 3.2 reviews the related literature. Section 3.3 describes the data and methodology. Section 3.4 documents forecast-definition differences and their implications for ICC levels, dispersion, and rankings. Section 3.5 presents measurement-error diagnostics and return-predictability tests, including heterogeneity across horizons and informational environments. Section 3.6

concludes.

3.2 Theoretical background

Analyst earnings forecasts play a central role in capital markets research because they summarize forward-looking expectations that enter valuation, trading decisions, and empirical proxies for market beliefs (e.g., Francis and Soffer (1997); Gebhardt et al. (2001)). A large literature studies analyst bias, accuracy, and revision behavior. In parallel, an asset-pricing and corporate finance literature relies on analyst forecasts as key inputs to ICC measures, which infer expected returns by equating observed prices to the discounted value of forecasted payoffs (e.g., Gebhardt et al. (2001); Gode and Mohanram (2003); Easton (2004); Li and Mohanram (2014)). Because ICCs solve for discount rates residually, the properties of forecast inputs are first-order determinants of the reliability of ICC-based expected return proxies. Concerns about forecast noise, implementation heterogeneity, and model sensitivity are therefore central in interpreting ICC evidence and in motivating diagnostic approaches that explicitly evaluate measurement error in expected-return proxies (e.g., Guay et al. (2011); Lee et al. (2021)).

A key feature of analyst data is that earnings expectations are not defined uniformly. Analysts and data vendors disseminate forecasts under alternative definitions, most prominently GAAP earnings and Street (non-GAAP) earnings. GAAP earnings follow standardized reporting rules, whereas Street earnings aim to reflect "core" performance by excluding selected items deemed transitory, non-operating, or otherwise non-recurring. Prior research documents systematic differences between GAAP and Street measures and emphasizes that the mapping between them varies across firms and over time (e.g., Bradshaw and Sloan (2002); Dechow et al. (2003); Landsman et al. (2007)). These definitional choices embed different judgments about persistence, the treatment of special items, and the appropriate long-run earnings base, implying that the same firm can exhibit meaningfully different expected earnings paths depending on whether forecasts are formed under GAAP or Street conventions.

Evidence on whether Street measures improve informational quality is mixed. Some event-study and disclosure-based evidence suggests that pro forma earnings can be more value relevant in certain settings, consistent with investors viewing exclusions as removing transitory noise (Bhattacharya et al. (2003)). Related work argues that

Street definitions can better capture persistent operating performance and may therefore command greater attention in pricing (Johnson and Schwartz Jr (2005)). At the same time, concerns remain that discretionary exclusions can be opportunistic, inconsistently applied, or tailored to managerial incentives, potentially reducing comparability and increasing noise (Dechow et al. (2003)). A further and critical point is that empirical conclusions can hinge on whether forecast definitions are properly aligned with realized benchmarks. Apparent superiority of one definition can reflect mechanical mismatches between the definition used to form expectations and the definition used to measure realizations rather than true informational dominance (e.g., Cohen et al. (2007); Bradshaw et al. (2018)). Consistent with this view, Bradshaw et al. (2018) show that explicit GAAP forecasts play a distinct role in investor interpretation and that analyses relying on indirect proxies for GAAP expectations can be contaminated by measurement error.

Recent research sharpens the identification issues in GAAP versus Street comparisons by emphasizing their structural linkage. Because non-GAAP earnings are constructed as GAAP earnings adjusted for exclusions, treating GAAP and non-GAAP as competing regressors in joint specifications can mechanically confound the accounting-level component with the adjustment component and lead to ambiguous interpretations. Christensen et al. (2025) formalize this point and show that a disciplined decomposition separating the accounting-consistent component from the adjustment yields clearer identification of whether investors price the GAAP level, the adjustment, or both. The implication for forward-looking settings is that Street forecasts are not a clean alternative to GAAP expectations: they embed GAAP-consistent expectations plus an adjustment. Designs that do not respect this structure risk conflating the two.

The same identification concern is particularly salient in ICC settings. ICCs map an entire forecasted earnings path into an implied discount rate through a valuation equation, so definitional differences in forecast inputs can translate into nontrivial differences in implied expected returns. This complicates simple Street–GAAP comparisons: it is unclear whether return-based evidence reflects GAAP-consistent expectations or the incremental adjustments embedded in Street forecasts. To keep this distinction explicit, the analysis re-parameterizes Street-based ICCs into a GAAP-based level and a Street–GAAP ICC differential and evaluates each part separately. This structure mirrors the identification logic in Christensen et al. (2025) while translating it into the ICC domain, and it allows tests to ask whether the definition-driven component contributes incremental information about expected returns beyond what is contained in GAAP-consistent expectations.

The economic significance of definition choice also depends on ICC model structure. Different ICC implementations load differently on near-term earnings levels, medium-run forecast horizons, and terminal value assumptions. Models that place greater weight on short-horizon payoffs are mechanically more sensitive to level differences across GAAP and Street forecasts, whereas models with heavier reliance on terminal values are more sensitive to differences in implied growth paths and long-run profitability. Consequently, definitional discrepancies need not translate uniformly into ICC differences across models; instead, their magnitude—and potentially their direction—can depend on how the valuation model maps forecast inputs into an implied discount rate. Understanding Street–GAAP differences therefore requires tracing how forecast levels and growth assumptions propagate through specific ICC structures rather than treating ICCs as a single homogeneous object.

This perspective also informs the mixed evidence in return-based tests using ICCs. While several studies report positive associations between ICCs and realized returns (e.g., Gode and Mohanram (2003); Hou et al. (2012); Li and Mohanram (2014)), other work finds weaker or less stable relations across samples and implementations (e.g., Guay et al. (2011); Easton (2004)). A prominent explanation is that ICCs contain economically meaningful discount-rate variation but are contaminated by measurement error arising from forecast noise and heterogeneous implementation choices, including horizon assumptions and terminal growth specifications. Recent work formalizes this view by evaluating expected-return proxies through their implied measurement-error variance (Lee et al. (2021)). Within this framework, forecast-definition choice matters because it alters the earnings path that ICC models translate into an implied discount rate. The relevant question is therefore empirical: whether ICCs constructed from GAAP versus Street forecasts exhibit different measurement-error properties and whether the definition-driven component of implied discount rates has incremental return relevance once the structural linkage between GAAP and Street measures is respected. This motivates evaluating GAAP- and Street-based ICCs side-by-side using both reliability diagnostics and predictive return regressions designed to respect their linkage.

Finally, the informational consequences of definitional differences are likely to be most pronounced when information frictions are higher. Disagreement and dispersion in beliefs are linked to uncertainty and limits to arbitrage (Diether et al. (2002)). Accounting complexity and intangible-intensive business models can increase discretion in classifying items as transitory versus core, potentially widening Street–GAAP gaps and making the adjustment component more contested. These considerations imply that forecast-definition choice should matter most in high-divergence states, among smaller or more opaque firms, and in sectors where performance measurement is less standardized. Taken together, the literature motivates a direct examination of forecast-definition choice within ICC estimation. If GAAP and Street forecasts embed systematically different earnings levels and growth paths, and if ICC models load differently on those components, then GAAP- and Street-based ICCs may diverge in levels, rankings, and reliability as expected-return proxies.

3.3 Data and methodology

3.3.1 Data and sample construction

The sample comprises firms in the S&P 500 with analyst earnings forecasts available in LSEG from 2004 to 2024. Focusing on large, widely followed firms improves the joint availability of Street and GAAP forecast definitions at multiple horizons. The final panel contains 125,754 firm-month observations from 902 unique firms.

For each firm-month t, I collect one- to five-year-ahead per-share earnings forecasts under two definitions: (i) Street earnings and (ii) GAAP earnings. To isolate definitional differences from timing differences, I require Street and GAAP forecasts to be observed for the same firm in the same month and align them by horizon. All cross-definition comparisons are therefore within firm-month and hold constant the information date. Table 3.1 summarizes sample composition and forecast availability across horizons.

Prices, market capitalization, and accounting inputs are obtained from LSEG, and I perform standard cross-checks against Compustat and I/B/E/S identifiers from WRDS and overlapping fields where available. All variables are denominated in U.S. dollars. Unless otherwise noted, all ICC estimates are winsorized at the 1st and 99th percentiles within month to reduce the influence of extreme realizations induced by model nonlinearity and occasional data errors.

Table 3.1: Sample and analyst forecast availability

Panel A: Sample Overview

Start Year	End Year	Unique Firms	Firm-Months
2004	2024	902	125,754

Panel B: Forecast Availability by Horizon

Horizon	Street Obs.	GAAP Obs.	Matched Obs.
FY1	121,813	119,041	118,908
FY2	121,975	119,929	119,897
FY3	121,752	118,107	118,082
FY4	116,954	112,572	112,445
FY5	114,781	110,397	110,229

Panel A reports the number of distinct firms and firm-month observations in the main sample of S&P 500 constituents from 2004–2024. Panel B reports firm-month availability of one- to five-year-ahead analyst earnings forecasts under the Street definition and under the GAAP definition, and the number of matched Street–GAAP observations within the same firm-month for each horizon ($FY1$–$FY5$). Forecasts are from I/B/E/S. Street forecasts use I/B/E/S *EPS* fields, while GAAP forecasts use I/B/E/S *GPS* (reported) fields.

3.3.2 Methodology

Implied cost of capital models

I estimate ICCs using five valuation-based models commonly applied in the expected-return literature: Gordon and Gordon (1997; GG), Claus and Thomas (2001; CT), Gebhardt, Lee, and Swaminathan (2001; GLS), Easton (2004; MPEG), and Ohlson and Juettner-Nauroth (2005; OJ). Consistent with prior work, I also report a composite measure to mitigate model-specific sensitivity and implementation noise (e.g., Hou et al. (2012); Li and Mohanram (2014); Gupta (2018); El Ghoul et al. (2018)).

For each firm-month and each model, I compute two ICCs that share identical prices and non-forecast inputs and differ only in the earnings-forecast definition used as input: ICC^{Street} uses Street forecasts and ICC^{GAAP} uses GAAP forecasts. To ensure economically meaningful discount-rate estimates, I exclude model-specific ICC solutions outside $[0, 1]$ prior to constructing the composite. The composite ICC is the simple average of available model-specific ICC estimates within each definition. Appendix B.1 details model implementations and input assumptions.

Empirical design

Because expected returns are unobservable, expected-return proxies must be evaluated indirectly. The empirical analysis proceeds in two steps. First, I assess the reliability of Street- and GAAP-based ICCs using the measurement-error variance (MEV) diagnostics of Lee et al. (2021). Second, I test whether GAAP-based ICCs and the Street–GAAP ICC differential predict subsequent realized returns in cross-sectional designs and in environments where forecast-definition divergence is high.

ICCs are nonlinear functions of forecast inputs and implementation assumptions. As a result, their cross-sectional and time-series variation may reflect both economically meaningful discount-rate variation and noise induced by forecast error or model sensitivity. Following Lee et al. (2021), I assess proxy reliability using MEV diagnostics computed separately in the cross-section and in the time series. The cross-sectional statistic evaluates whether an ICC proxy captures cross-sectional variation in expected returns within a month, while the time-series statistic evaluates whether the proxy captures within-firm variation in expected returns over time. In both dimensions, lower scaled measurement-error variance indicates a higher signal-to-noise ratio (i.e., stronger comovement with subsequent realized returns relative to the proxy's own variability). I compute these diagnostics for ICC^{Street} and ICC^{GAAP} for each model and for the composite, and use them to benchmark the interpretation of subsequent return tests. Appendix B.2 provides formal definitions and implementation details.

To isolate the economic implications of forecast-definition choice, I summarize definition effects in implied discount-rate space using the Street–GAAP ICC differential,

$$ICC_{i,t}^{Diff} \equiv ICC_{i,t}^{Street} - ICC_{i,t}^{GAAP}.$$

I then estimate pooled panel regressions that relate future realized returns to $ICC_{i,t}^{GAAP}$ and $ICC_{i,t}^{Diff}$:

$$r_{i,t\to t+h} = \alpha + \beta_1\, ICC_{i,t}^{GAAP} + \beta_2\, ICC_{i,t}^{Diff} + \gamma' X_{i,t} + \mu_{\text{Ind}} + \lambda_t + \varepsilon_{i,t+h}, \tag{3.1}$$

where $r_{i,t\to t+h}$ is the annualized total return over horizon $h \in \{12, 24, 36\}$ months, $X_{i,t}$ includes log market capitalization and book-to-market, and μ_{Ind} and λ_t denote industry and time fixed effects. The fixed effects absorb persistent differences across industries and common time variation in expected returns, so identification comes from within-industry, within-time variation in ICC measures. For comparability with the ICC literature, I also report single-proxy specifications and joint specifications that include ICC^{Street} and ICC^{GAAP} simultaneously.

This empirical design is closely related to the canonical cross-sectional return-predictability framework in which slope coefficients are estimated from repeated period-by-period cross-sections (Fama, 1973). The pooled specification is convenient in large firm-time panels and allows a unified treatment of controls and fixed effects; inference must account for residual dependence inherent to panel return data. Following standard guidance for finance panel regressions, I report heteroskedasticity-robust standard errors clustered at the firm level (Petersen, 2009). The inclusion of size and book-to-market follows classic evidence on size- and value-related patterns in the cross section of stock returns (Banz, 1981; Fama and French, 1995). Finally, the use of pooled panel regressions with industry and time fixed effects mirrors recent disagreement-based designs in adjacent settings, which emphasize absorbing industry-level and time-varying components when disagreement measures vary systematically across sectors (Gibson Brandon et al., 2021).

I re-estimate (3.1) in states where Street and GAAP expectations diverge most and in subsamples commonly associated with greater informational frictions. Specifically, *HighDiv* is an indicator equal to one for firm-months in the top quintile (within month) of the average absolute Street–GAAP forecast divergence across selected horizons, and zero otherwise. I further examine heterogeneity across firm-size terciles and across industry groupings. These tests assess whether GAAP-based ICCs contain incremental return information precisely when forecast-definition differences are largest.

The two steps are interpreted jointly. MEV diagnostics speak directly to proxy reliability and provide a benchmark for whether cross-definition differences plausibly reflect signal versus noise. The regression results then quantify the economic relevance of ICC^{GAAP} and of the Street–GAAP ICC differential.

3.4 Discrepancies in earnings forecasts and ICCs

This section traces how earnings-definition differences in analyst forecasts propagate into ICC estimates. The analysis proceeds from forecast inputs to ICC outputs. First, I document systematic differences between Street and GAAP forecast levels and implied growth across horizons. Second, I map these differences into model-specific Street–GAAP ICC differentials and characterize their cross-sectional dispersion and time variation. Third, I examine the implications for empirical designs that rely on cross-sectional ranks by quantifying ranking stability and tail reclassification when moving from Street to

GAAP inputs.

3.4.1 Forecast differences: levels and implied growth

Table 3.2 reports summary statistics for analyst earnings forecasts per share under Street and GAAP definitions across horizons $FY1$ to $FY5$. The statistics are computed on matched firm-month observations (same firm and month under both definitions) and summarize the cross-sectional distribution within each horizon (mean, median, dispersion, and tail percentiles).

Table 3.2 shows that Street and GAAP forecasts differ systematically in levels. At the one-year horizon ($FY1$), the mean Street forecast equals 5.2 compared to 4.3 for GAAP; median forecasts are 3.1 and 2.8, respectively. Level differences persist across horizons, although the gap narrows somewhat at longer maturities. Dispersion also differs meaningfully: GAAP forecasts exhibit substantially higher cross-sectional standard deviations at short horizons (e.g., 32.8 versus 19.2 at $FY1$). This pattern is consistent with Street definitions filtering out selected items, which mechanically reduces cross-sectional variation in the forecasted earnings measure.

Table 3.2: Summary statistics of Street and GAAP earnings forecasts

		Street Forecasts					GAAP Forecasts				
Horizon	*N*	Mean	Median	Std	P1	P99	Mean	Median	Std	P1	P99
FY1	118,908	5.199	3.060	19.222	-1.440	30.747	4.345	2.840	32.799	-3.580	28.984
FY2	119,897	5.868	3.425	20.440	-0.085	33.671	5.617	3.272	20.495	-0.220	33.116
FY3	118,082	6.645	3.850	22.939	0.175	39.000	6.484	3.740	26.315	0.107	38.817
FY4	112,445	7.514	4.284	26.083	0.257	45.100	7.447	4.235	30.184	0.180	44.530
FY5	110,229	8.336	4.709	29.384	0.300	50.830	8.471	4.720	34.988	0.228	52.131

This table reports summary statistics of analyst earnings forecasts (per share) under the Street and GAAP definitions for forecast horizons $FY1$–$FY5$. Statistics are computed on matched Street–GAAP observations within each horizon (same firm and month). Forecasts are from I/B/E/S. Street forecasts correspond to I/B/E/S *EPS* forecast fields; GAAP forecasts correspond to I/B/E/S *GPS* (reported) forecast fields. N denotes the number of matched observations used for the horizon. P1 and P99 denote the 1st and 99th percentiles of the cross-sectional distribution within the horizon.

Table 3.3 reports the corresponding annualized implied growth rates constructed from forecasts relative to last realized earnings under each definition. The table summarizes the cross-sectional distribution of these growth rates within each horizon (mean, median, dispersion, and interdecile range), using matched observations. Table 3.4 provides a more detailed view of growth by breaking it down into sequential year-over-year steps between adjacent forecast horizons.

Together, the two tables reveal a systematic level–growth offset. Although Street

Table 3.3: Implied growth in analyst's forecast under Street and GAAP definitions

		Street Growth (in %)					GAAP Growth (in %)				
Horizon	N	Mean	Median	Std	P10	P90	Mean	Median	Std	P10	P90
FY1	104,022	7.97	8.45	28.95	-18.60	31.87	12.20	10.13	43.35	-31.43	56.65
FY2	108,125	10.47	9.66	18.77	-5.31	25.29	15.63	11.73	28.29	-9.40	42.92
FY3	107,384	10.82	9.92	14.23	-1.00	22.24	14.97	11.79	21.55	-3.03	34.70
FY4	102,496	10.90	10.01	11.79	1.11	20.86	14.56	11.89	17.75	-0.17	31.06
FY5	100,647	10.75	9.91	10.21	2.13	19.99	14.18	11.75	15.61	1.23	29.17

This table reports annualized growth rates constructed from analyst earnings forecasts under the Street and GAAP definitions. For horizon FYi, growth is computed as $(F_i/Act)^{1/i} - 1$, where Act denotes last realized earnings per share and F_i represents the analysts' forecast under the corresponding definition. The sample is restricted to observations with positive last realized earnings under each definition. Forecasts are from I/B/E/S. Street forecasts correspond to I/B/E/S *EPS* forecast fields; GAAP forecasts correspond to I/B/E/S *GPS* (reported) forecast fields. Percentiles P10 and P90 refer to the cross-sectional distribution within each horizon.

forecasts are generally higher in levels, GAAP forecasts consistently imply steeper growth rates. At $FY1$, for example, median implied growth is 10.13% under GAAP, compared with 8.45% under Street, and this ordering persists across all subsequent horizons.

Table 3.4: Sequential forecast growth rates: EPS vs. GAAP

		EPS Growth (in %)		GAAP Growth (in %)	
Horizon	N	Mean	Median	Mean	Median
FY0 → FY1	97,888	14.96	8.90	42.07	11.41
FY1 → FY2	97,888	16.37	11.07	29.65	12.72
FY2 → FY3	97,888	12.20	10.78	17.48	11.55
FY3 → FY4	97,888	11.81	10.23	13.64	11.00
FY4 → FY5	97,888	10.78	9.59	13.01	10.47

Notes: The table reports sequential year-over-year forecast growth rates for EPS and GAAP earnings. For FY1, growth is computed as $(FY1/Act) - 1$, where Act denotes last realized earnings. For later horizons, growth is computed as $(F_t/F_{t-1}) - 1$. The sample is restricted to observations with positive denominator and numerator values for the respective growth calculation.

Table 3.4 helps localize where this difference arises. The contrast is strongest in the transition from $L0$ to $F1$, where mean GAAP growth reaches 42.1%, compared with 15.0% under Street. A plausible interpretation is that realized GAAP earnings in the base period are more strongly affected by transitory charges such as impairments, restructuring expenses, or write-downs, whereas forward forecasts reflect expectations that these items will not persist and that earnings will revert toward more representative levels (Doyle et al. (2003); Kolev et al. (2008)). Because Street earnings exclude more of these items, they start from a higher and smoother base, which mechanically implies flatter growth. GAAP forecasts, in contrast, begin from a lower base and therefore imply a stronger near-term recovery, even when their forecasted levels remain below Street levels.

The sequential evidence also shows that the difference does not disappear immediately. From $F1$ to $F2$, mean implied growth remains substantially higher under GAAP than under Street (29.7% versus 16.4%), indicating that the Street–GAAP level difference continues to shape near-term forecast dynamics beyond the first step. At the same time, the gap narrows steadily at longer horizons, falling to 13.6% versus 11.8% from $F3$ to $F4$ and to 13.0% versus 10.8% from $F4$ to $F5$. This fading pattern is consistent with temporary Street–GAAP adjustments mattering most for near-term reported earnings and becoming less important as forecasts move toward a more common medium-run earnings path.

This level–growth offset has direct implications for ICC estimation. Because ICC models differ in the weight they place on near-term growth, forecast levels, and long-run anchors, the forecast-definition differences will not map into a uniform Street-GAAP ICC differential across models.

3.4.2 From forecast differences to ICC differences

Table 3.5 summarizes the level of implied costs of capital computed from Street versus GAAP forecast inputs. For each ICC model (and the composite), the table reports the cross-sectional distribution of ICC^{Street} and ICC^{GAAP} across firm-months, expressed in percentage points. This benchmark is useful for two reasons: it situates the magnitude of the estimates relative to the ICC literature, and it shows whether definition choice shifts the level of implied discount rates on average.

The table indicates that average ICC magnitudes vary across models, consistent with known structural differences in how valuation formulas load on near-term cash flows versus terminal components. For the composite measure, the mean ICC equals 9.13% under Street inputs and 9.41% under GAAP inputs. Differences are more pronounced in growth-sensitive models: for example, MPEG yields a mean ICC of 10.64% under Street and 11.55% under GAAP, whereas GG exhibits smaller average differences (6.65% versus 6.25%). Although these average shifts may appear modest, a one-percentage-point change in an implied discount rate is economically meaningful relative to typical estimates of the annual equity risk premium reported in survey-based benchmarks (e.g., around 6–7% in Welch, 2000).

Figure 3.1 maps forecast differences into Street–GAAP ICC differentials across valuation models. The figure implements a sorting design: firms are assigned to deciles

Table 3.5: Implied cost of capital levels by valuation model: Street versus GAAP forecast inputs

		Street-based ICC (%)					GAAP-based ICC (%)				
Model	N	Mean	Median	Std.	P10	P90	Mean	Median	Std	P10	P90
CT	115,560	7.85	7.48	3.33	4.30	11.60	7.81	7.37	3.53	4.07	11.75
GLS	115,452	9.90	9.70	3.66	5.54	14.21	9.72	9.52	3.67	5.33	14.07
MPEG	104,293	10.64	9.43	5.28	6.44	15.78	11.55	9.89	6.53	6.44	18.11
OJ	104,068	10.63	9.45	5.28	6.39	15.79	11.54	9.91	6.53	6.41	18.08
GG	113,201	6.65	5.98	3.72	3.06	10.72	6.25	5.62	3.74	2.56	10.32
Composite	118,360	9.13	8.49	3.80	5.65	13.06	9.41	8.61	4.33	5.56	13.74

This table reports implied costs of capital (ICCs) estimated using five valuation models: CT, GLS, MPEG, OJ, and GG, and their composite (simple average of available model estimates). For each model, ICCs are computed twice using identical prices, book values, and model assumptions, differing only in whether forecast inputs are taken under the Street or GAAP earnings definition. Forecast inputs are from I/B/E/S: Street forecasts correspond to I/B/E/S *EPS* forecast fields; GAAP forecasts correspond to I/B/E/S *GPS* (reported) forecast fields. Values are expressed in percent. ICCs are winsorized at the 1st and 99th percentiles within each month. N denotes the number of matched Street–GAAP observations available for the given model.

based on the Street–GAAP forecast difference in the relevant forecast input, and the bars report the median ICC^{Diff}, within each decile (in percentage points). Panel A sorts on differences in forecast *levels*, while Panel B sorts on differences in implied forecast *growth*. In both panels, decile 1 corresponds to observations where GAAP exceeds Street the most (a strongly negative Street–GAAP forecast difference), and decile 10 corresponds to observations where Street exceeds GAAP the most.

Two patterns emerge. First, the mapping from forecast differences to ICC^{Diff} is strongly model dependent, consistent with valuation structures placing different weights on near-term earnings versus long-run growth. In Panel A, models with greater sensitivity to near-term earnings translate level differences into largely monotonic movements in ICC^{Diff}. When Street levels substantially exceed GAAP (upper deciles), ICC^{Diff} is typically negative, consistent with higher forecasted near-term payoffs requiring a higher discount rate to match the same price.

Second, Panel B shows that differences in implied growth generate the largest GAAP–Street ICC differentials, particularly in the left tail of the growth-difference distribution. For growth-sensitive models such as MPEG and OJ, ICC^{Diff} is most negative in decile 1, where GAAP-implied growth substantially exceeds Street-implied growth. In these observations, the median ICC^{Diff} reaches magnitudes in excess of three percentage points, implying that ICC^{GAAP} can be markedly higher than ICC^{Street} when GAAP forecasts embed much steeper growth trajectories. Moving across deciles toward cases where Street-implied growth exceeds GAAP-implied growth (upper deciles), ICC^{Diff}

turns positive, but the magnitude is generally smaller than in the left tail. Thus, the figure highlights an economically important asymmetry: the largest GAAP–Street ICC differentials arise when GAAP growth dominates Street growth, not when Street growth dominates GAAP growth.

Overall, Figure 3.1 shows that forecast-definition effects on ICCs are not uniform shifts. Instead, they arise from predictable interactions between (i) whether Street–GAAP differences appear primarily in forecast levels or in implied growth and (ii) how each ICC model loads on those components, with the most pronounced effects occurring in the tails of the forecast-difference distribution.

The sorting evidence clarifies the mapping from forecast differences to ICC differences, but it does not reveal how large and persistent these GAAP–Street ICC differentials are in the cross-section and over time. Figure 3.2 addresses this by plotting, for each calendar year, the cross-sectional fractiles of the composite Street–GAAP ICC differential. The figure summarizes how the center and tails of the difference evolve over time.

The median ICC^{Diff} is close to zero in many years, consistent with the modest average differences in Table 3.5. At the same time, dispersion remains economically meaningful: in several periods the distance between the 10th and 90th percentiles spans multiple percentage points, indicating that a sizeable subset of firms experiences large positive or negative GAAP–Street ICC differentials even when the center of the distribution is stable. Two implications follow. First, because the composite ICC averages across valuation structures with different sensitivities to forecast levels and implied growth, persistent dispersion in the composite ICC^{Diff} indicates that forecast-definition effects are not confined to a single model class; they survive aggregation. Second, the repeated appearance of wide tail differences suggests that forecast-definition effects are not isolated episodes but instead reflect recurring, state-dependent separation between ICC^{Street} and ICC^{GAAP}.

Table 3.6 complements the time-series depiction by summarizing the cross-sectional distribution of ICC^{Diff} separately by ICC model. For each model, the table reports central tendency and dispersion (including tail percentiles), allowing the magnitude and heterogeneity of forecast-definition effects to be compared across valuation structures. Mean GAAP–Street ICC differentials are close to zero for some models (e.g., 0.04 p.p. for CT), but materially negative for growth-sensitive models (e.g., –0.92 p.p for MPEG and –0.91 p.p for OJ). Dispersion is substantial: the standard deviation exceeds 3.6 percentage

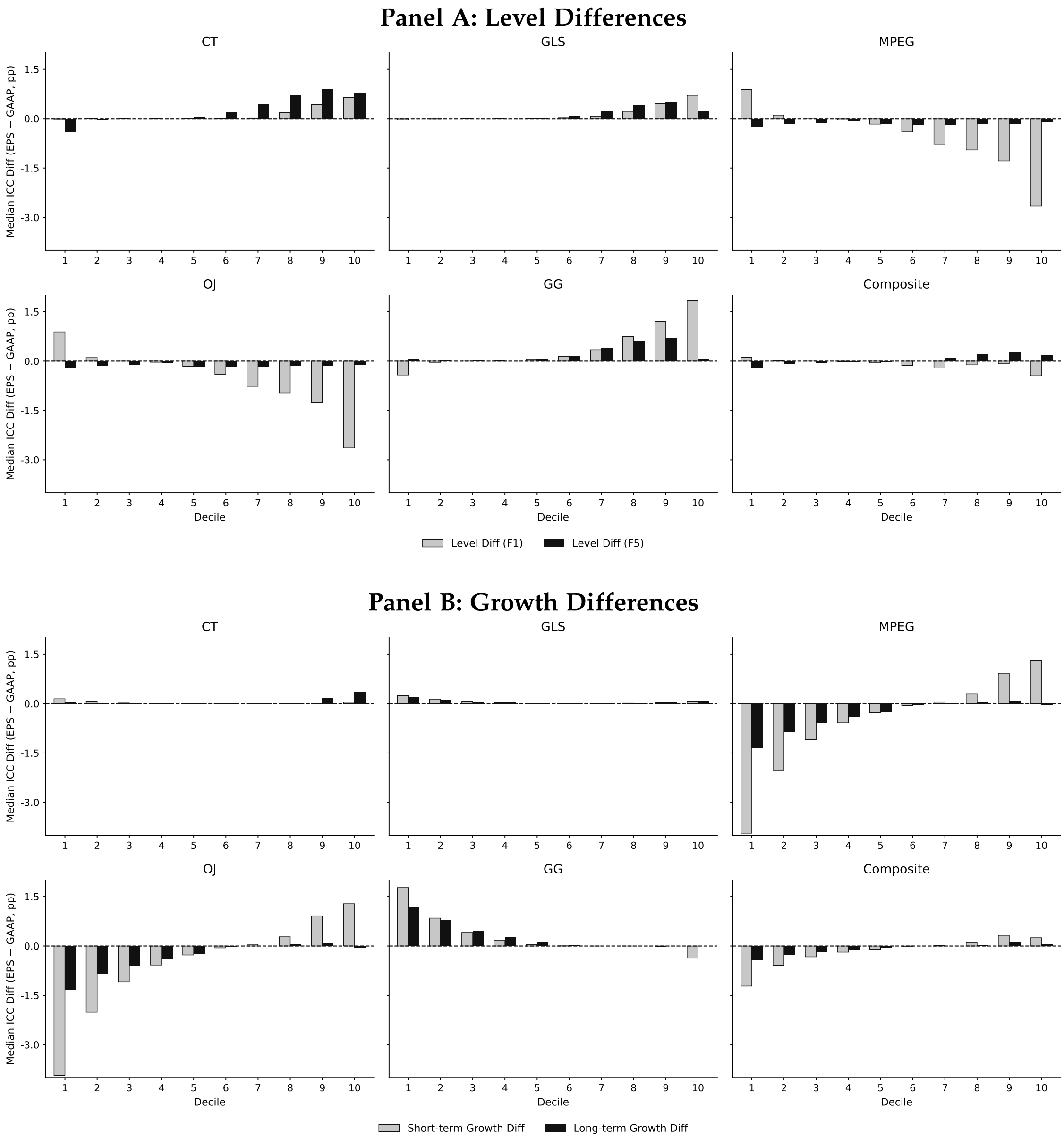


Figure 3.1: ICC differences by level and growth divergence deciles

This figure examines how forecast-definition differences between Street and GAAP earnings map into ICC differences, $ICC^{Diff} \equiv ICC^{Street} - ICC^{GAAP}$. Panel A plots the median ICC^{Diff} (in percentage points) across deciles formed on forecast *level* differences (Street minus GAAP), measured separately for one-year-ahead (FY1) and five-year-ahead (FY5) forecasts. Panel B plots the median ICC^{Diff} across deciles formed on differences in *implied growth* (Street minus GAAP). Short-term growth is defined as the annualized growth implied by forecasts from year 1 to year 2 (FY1→FY2), and long-term growth is defined analogously from year 1 to year 5 (FY1→FY5). In both panels, decile 1 corresponds to observations where GAAP exceeds Street the most (a strongly negative Street–GAAP forecast difference), and decile 10 corresponds to observations where Street exceeds GAAP the most. ICCs are estimated separately under Street and GAAP inputs using GG, GLS, CT, MPEG, and OJ, and then differenced to obtain ICC^{Diff}. The sample consists of matched Street and GAAP analyst forecasts for S&P 500 firms from 2004–2024.

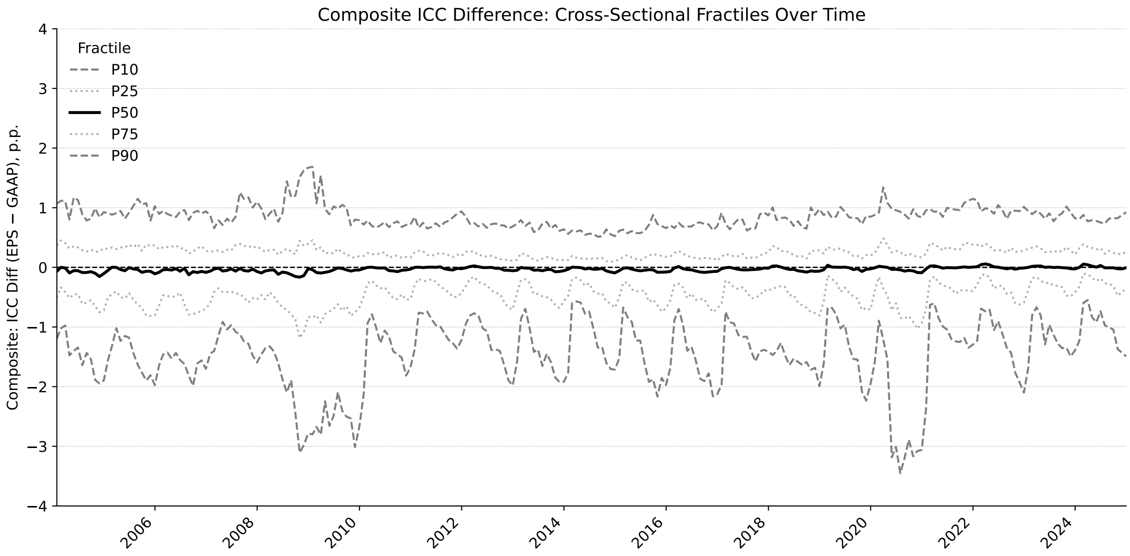


Figure 3.2: Distribution of composite ICC differences from 2004 to 2024

This figure plots the annual cross-sectional 10th, 25th, 50th, 75th, and 90th percentiles of the composite GAAP–Street ICC differential, $ICC^{Diff} \equiv ICC^{Street} - ICC^{GAAP}$. ICC^{Street} is estimated using Street (analyst-adjusted) earnings forecasts (EPS), whereas ICC^{GAAP} is estimated using reported-earnings forecasts (GPS). The composite ICC is computed as the simple average of five valuation-model estimates: GG (Gordon and Gordon, 1997), GLS (Gebhardt et al., 2001), CT (Claus and Thomas, 2001), MPEG (Easton, 2004), and OJ (Ohlson and Juettner-Nauroth, 2005). ICCs are estimated using matched Street and GAAP analyst forecasts for S&P 500 firms over 2004–2024.

points for MPEG and OJ, and the 10th percentile reaches –3.62 percentage points for MPEG. Even the composite GAAP–Street ICC differential exhibits a standard deviation of 1.92 percentage points. Overall, the table shows that while average forecast-definition effects on ICCs can be small, the cross-sectional variation in the Street–GAAP ICC differential is economically meaningful.

Table 3.6: Sensitivity of ICC models to earnings choice: Street–GAAP ICC difference

Model	*N*	Mean (%)	Median (%)	Std (%)	P10 (%)	P90 (%)
CT	115,560	0.04	0.00	1.53	-0.97	1.21
GLS	115,452	0.17	0.03	0.70	-0.18	0.80
MPEG	104,293	-0.92	-0.14	3.61	-3.62	1.09
OJ	104,068	-0.91	-0.14	3.60	-3.60	1.09
GG	113,201	0.40	0.05	1.43	-0.15	1.62
Composite	118,360	-0.28	-0.02	1.92	-1.45	0.85

This table reports the ICC difference, defined as Street-based minus GAAP-based ICC, for each valuation model and for the composite ICC. Values are expressed in percentage points. ICCs are constructed using forecast inputs from I/B/E/S: Street forecasts correspond to I/B/E/S *EPS* forecast fields; GAAP forecasts correspond to I/B/E/S *GPS* (reported) forecast fields. For each model, the difference is computed on the matched Street–GAAP sample used to construct the model-specific ICCs.

3.4.3 Implications for ICC rankings

Table 3.7 summarizes how closely Street- and GAAP-based ICC constructions align in the cross-section. For each valuation model (and the composite), the table reports month-by-month Pearson and Spearman rank correlations between ICC^{Street} and ICC^{GAAP} averaged over time, along with a tail-turnover metric that quantifies decile reclassification when moving from Street to GAAP inputs.

Overall alignment is strong. Pearson correlations range from 0.834 (MPEG and OJ) to 0.982 (GLS), with Spearman correlations of comparable magnitude, indicating substantial comovement in broad cross-sectional ranks.

Table 3.7: Rank correlations and tail turnover of Street- and GAAP-based ICCs

Model	*N*	Pearson	Spearman	Top 10% turnover (in %)	Bottom 10% turnover (in %)
CT	115,560	0.903	0.909	22.7	17.9
GLS	115,452	0.982	0.984	8.1	7.9
MPEG	104,293	0.834	0.826	27.7	33.2
OJ	104,068	0.834	0.829	27.7	31.8
GG	113,201	0.927	0.917	15.5	25.8
Composite	118,360	0.896	0.922	21.6	15.8

This table reports rank similarity between Street- and GAAP-based ICC estimates within each valuation model. Pearson and Spearman correlations are computed across firms within each month and then averaged over months (using the matched Street–GAAP sample available for the model). Top (bottom) 10% turnover denotes the fraction of firms classified in the top (bottom) Street ICC decile that are not in the corresponding GAAP ICC decile within the same month. ICCs are constructed using forecast inputs from I/B/E/S: Street forecasts correspond to I/B/E/S *EPS* forecast fields; GAAP forecasts correspond to I/B/E/S *GPS* (reported) forecast fields.

High rank correlations, however, can mask economically relevant differences in the tails, where many empirical designs concentrate identification (e.g., portfolio sorts and long–short spreads). Tail turnover in Table 3.7 shows meaningful instability in these regions. For example, 27.7% of firms classified into the top decile under ICC^{Street} do not remain in the top decile under ICC^{GAAP} for MPEG (and similarly for OJ). Even for the composite ICC, top-decile turnover is 21.6%, and bottom-decile turnover is 15.8%.

These reclassification rates imply that forecast-definition choice can materially alter the composition of "high-ICC" and "low-ICC" portfolios and, more generally, can change which firms are treated as being in the extremes of the expected-return-proxy distribution. Thus, small average differences and high overall correlations can coexist with substantial instability exactly where many ICC applications draw their strongest contrasts. This evidence is consistent with the wide tail dispersion in the composite GAAP–Street ICC differential documented in Figure 3.2.

Taken together, the results indicate that forecast-definition choice affects not only ICC levels but also cross-sectional ordering in a way that is concentrated in the tails and varies across valuation models. The next section examines whether these Street–GAAP ICC differentials are return relevant across horizons and informational environments.

3.5 Reliability and performance of ICCs

This section examines whether ICCs constructed from Street versus GAAP forecasts differ in (i) their measurement-error properties and (ii) their relevance for subsequent realized returns. The analysis proceeds in two steps. First, scaled measurement-error variance diagnostics following Lee et al. (2021) assess whether forecast-definition choice materially affects the noise embedded in ICC proxies. Second, predictive return regressions evaluate whether ICC^{GAAP} and ICC^{Diff} contain incremental return-relevant information.

3.5.1 Measurement-error variance of Street- and GAAP-based ICCs

Table 3.8 reports scaled measurement-error variance (SVar) diagnostics following Lee et al. (2021) for Street- and GAAP-based ICC constructions. The diagnostics provide an explicit reliability benchmark for expected-return proxies, evaluated separately in the time series and in the cross-section. Appendix B.2 defines SVar; throughout, more negative values indicate stronger comovement with subsequent realized returns relative to the proxy's unconditional variability.

Across all models and both forecast definitions, the time-series SVar estimates are negative and statistically significant. For the composite ICC, the time-series SVar equals -0.0021 under Street inputs and -0.0026 under GAAP inputs, with similar patterns for CT, GLS, MPEG, OJ, and GG. This evidence indicates that both ICC^{Street} and ICC^{GAAP} behave as admissible expected-return proxies in the time-series dimension.

Importantly, paired difference tests do not reveal systematic time-series reliability advantages for either definition. The composite difference is small (0.0005) and statistically insignificant, and the same conclusion holds across the individual models. Thus, there is no consistent evidence that one forecast definition yields uniformly lower time-series measurement error.

In the cross-section, SVar magnitudes are economically small and generally statistically insignificant for both definitions. For the composite ICC, cross-sectional SVar equals -0.0004 under Street inputs and -0.0002 under GAAP inputs, both indistinguishable from zero. While some model-level difference tests are statistically significant (GG, MPEG, and OJ), the effects are modest in magnitude and not systematic across models. Overall, the cross-sectional diagnostics do not indicate a robust reliability advantage for either earnings definition.

Table 3.8: Measurement-error variance by ICC model: Street versus GAAP

ICC Model	Earnings Type	Time-Series SVar	Cross-Sectional SVar
CT	Street	-0.0014***	-0.0003
	GAAP	-0.0015***	-0.0002
	Difference (Street−GAAP)	0.0002	-0.0001
GLS	Street	-0.0012***	0.0002
	GAAP	-0.0012***	0.0003
	Difference (Street−GAAP)	0.0000	-0.0000
MPEG	Street	-0.0024***	0.0000
	GAAP	-0.0032***	0.0009
	Difference (Street−GAAP)	0.0010	-0.0009**
OJ	Street	-0.0023***	0.0000
	GAAP	-0.0030***	0.0009
	Difference (Street−GAAP)	0.0010	-0.0009**
GG	Street	-0.0012***	0.0001
	GAAP	-0.0017***	-0.0002
	Difference (Street−GAAP)	0.0004	0.0003*
Composite	Street	-0.0021***	-0.0004
	GAAP	-0.0026***	-0.0002
	Difference (Street−GAAP)	0.0005	-0.0002

This table reports scaled measurement-error variances following Lee et al. (2021). Forecast inputs are from I/B/E/S: Street forecasts correspond to I/B/E/S *EPS* forecast fields; GAAP forecasts correspond to I/B/E/S *GPS* (reported/GAAP) forecast fields. For each valuation model and for the Composite ICC, SVar is computed as $\mathrm{Var}(\widehat{ICC}) - 2\,\mathrm{Cov}(r_{t+1}, \widehat{ICC})$ using one-month-ahead total returns; reported values are multiplied by 100. For Street and GAAP rows, stars denote significance from a t-test of SVar against zero. For the Difference row (Street−GAAP), stars denote significance from a paired test of SVar objects (paired across firms for time-series SVar and across months for cross-sectional SVar). ***, **, and * denote 1%, 5%, and 10% levels.

Taken together, the MEV diagnostics validate both Street- and GAAP-based ICC constructions as admissible expected-return proxies in the time-series. At the same time, the evidence does not indicate that one forecast definition yields systematically lower measurement error than the other at an aggregate level.

These results motivate a shift in emphasis. Because the two ICC constructions exhibit broadly comparable aggregate reliability and are mechanically linked through their forecast inputs, the key question is whether the adjustment embedded in Street forecasts contributes incremental return-relevant variation beyond GAAP-consistent expectations. The subsequent analysis therefore re-parameterizes ICC^{Street} into ICC^{GAAP} and ICC^{Diff} and evaluates the return relevance of each across horizons and informational environments.

3.5.2 Street versus GAAP ICC

Table 3.9 reports baseline return regressions that relate ICCs constructed from Street versus GAAP analyst forecast inputs to subsequent realized returns. The dependent variable is the next-12-month annual total return. All specifications include the same controls and industry and time fixed effects, and standard errors are clustered by firm.

Columns (1) and (2) consider each composite ICC measure separately. The composite ICC^{Street} is positively associated with next-year returns (coefficient: 0.928), and the composite ICC^{GAAP} is also positively associated with next-year returns (coefficient: 0.800); both estimates are statistically significant. Consistent with ICCs serving as forward-looking expected-return proxies, each construction exhibits positive return relevance when examined in isolation.

Column (3) includes both ICC^{Street} and ICC^{GAAP} in the same regression. Because the two measures share prices and most non-forecast inputs and differ only in the earnings-definition content of forecast inputs, this joint specification assesses whether either definition contributes incremental return-relevant variation beyond what is common to both. In the joint regression, the Street-based measure remains strongly positive (0.719), while the incremental coefficient on the GAAP-based measure is smaller and statistically weak (0.232). At the 12-month horizon, the shared component between the two constructions appears to account for most of the return-relevant variation captured by the joint specification.

The next set of tests makes the structural linkage between forecast definitions explicit by re-parameterizing the Street-based ICC into a GAAP-based level and a Street–GAAP ICC differential. This design evaluates whether the definition-driven component embedded in Street forecasts is return relevant beyond GAAP-consistent expectations, and whether these relations vary across horizons and informational environments.

Table 3.9: Baseline return regressions: Street versus GAAP composite ICC

	(1) Street only	(2) GAAP only	(3) Joint
ICC^{Street}	0.928***		0.719***
	(5.29)		(2.63)
ICC^{GAAP}		0.800***	0.232
		(5.27)	(1.03)
Log(Size)	-0.008**	-0.008**	-0.007*
	(-2.12)	(-2.08)	(-1.95)
Book-to-Market	-0.000	-0.000**	-0.000***
	(-1.07)	(-2.50)	(-2.58)
Observations	97,239	95,371	95,254
Adj. R^2	0.177	0.174	0.174

The dependent variable is the firm's subsequent 12-month buy-and-hold annual total return (in decimals), measured from month $t+1$ to $t+12$ relative to the month-t ICC estimate. ICC^{Street} and ICC^{GAAP} are composite implied costs of capital constructed as the simple average of five valuation-model ICC estimates (GG, GLS, CT, MPEG, and OJ), computed using Street versus GAAP analyst forecast inputs, respectively. Forecast inputs are from I/B/E/S: Street forecasts correspond to I/B/E/S *EPS* forecast fields; GAAP forecasts correspond to I/B/E/S *GPS* (reported/GAAP) forecast fields. All specifications include industry (GICS sector) fixed effects and time fixed effects, and control for size (log market capitalization) and book-to-market, each measured at month t. Standard errors are heteroskedasticity-robust and clustered at the firm level. ***, **, and * denote significance at the 1%, 5%, and 10% levels.

3.5.3 GAAP ICC and the Street–GAAP ICC differential

Table 3.10 reports predictive return regressions that separate the Street-based composite ICC into (i) a GAAP-based level and (ii) a Street–GAAP differential that captures the definition-driven component embedded in Street forecasts. The table examines how the return relevance of these two terms varies with the return horizon. Each column reports a separate regression using annualized realized returns over 12, 24, and 36 months; all specifications include the same controls and industry and time fixed effects as in the baseline, with standard errors clustered by firm.

At the 12-month horizon, both terms are positively associated with subsequent returns and statistically significant. The coefficient on the GAAP-based ICC is 0.951, and the coefficient on the Street–GAAP ICC differential is 0.719. Thus, in the short run, realized returns covary with both the GAAP-consistent component of implied discount rates and the definition-driven component embedded in Street-based ICCs.

At longer horizons, the associations attenuate and become less precisely estimated. At 24 months, the GAAP-based ICC remains positive (0.226) but is smaller in magnitude, while the coefficient on the Street–GAAP ICC differential is not statistically distinguishable from zero. At 36 months, neither term shows a reliable association with subsequent realized returns in this specification.

Table 3.10: Return regressions with GAAP ICC and the Street–GAAP ICC differential across horizons

	Horizon (annualized)		
	12m	24m	36m
ICC^{GAAP}	0.951***	0.226**	-0.044
	(5.24)	(2.27)	(-0.54)
ICC^{Diff}	0.719***	0.154	0.027
	(2.63)	(0.87)	(0.20)
Observations	95,254	94,752	94,251
Adj. R^2	0.174	0.186	0.178

The dependent variable is the firm's subsequent buy-and-hold annual total return over the stated horizon, expressed in annualized terms (12m: raw 12-month return; 24m and 36m: multi-year cumulative returns annualized to a per-year rate). All regressors are measured at month t, and returns are measured from $t+1$ onward to avoid look-ahead bias. ICC^{GAAP} is the composite implied cost of capital computed as the simple average of five valuation-model ICC estimates (GG, GLS, CT, MPEG, and OJ) using GAAP analyst forecast inputs. The Street–GAAP ICC differential is defined as $ICC_{i,t}^{Diff} \equiv ICC_{i,t}^{Street} - ICC_{i,t}^{GAAP}$ for the same firm-month, where ICC^{Street} uses Street earnings forecasts and ICC^{GAAP} uses GAAP forecasts. Forecast inputs are from I/B/E/S: Street forecasts correspond to I/B/E/S *EPS* forecast fields; GAAP forecasts correspond to I/B/E/S *GPS* (reported/GAAP) forecast fields. Each specification includes controls for size (log market capitalization) and book-to-market (measured at t), as well as industry (GICS sector) and time fixed effects. Standard errors are heteroskedasticity-robust and clustered at the firm level. ***, **, and * denote significance at the 1%, 5%, and 10% levels.

3.5.4 High earnings-divergence states

Section 3.4 shows that the Street–GAAP gap in analyst expectations varies widely across firm-months. If GAAP-based expectations are most informative when the two forecast definitions separate most, then the return association of the GAAP-based ICC component should be stronger in high-divergence states.

Table 3.11 evaluates this idea using an interaction design that compares firm-months in the top quintile of forecast-definition divergence (*HighDiv*) to all remaining observations. The table reports implied marginal effects of ICC^{GAAP} and ICC^{Diff} on annualized realized returns over 12, 24, and 36 months. Low-divergence effects correspond to the main coefficients. High-divergence effects are computed as the sum of the main and interaction terms and are evaluated using Wald tests.

Three patterns emerge. First, when divergence is low, both the GAAP-based ICC and the Street–GAAP ICC differential are positively associated with subsequent returns at shorter horizons: at 12 months, the coefficients are 0.720 and 0.898 (both statistically significant), and both remain positive and statistically significant at 24 months. Second, when divergence is high, the implied ICC^{GAAP} effect is materially larger at 12 months (1.347) and remains positive at 24 months (0.319), while it is close to zero by 36 months. Third, the Street–GAAP ICC differential is positive under high divergence at 12 months (0.944) but becomes smaller and statistically weak at longer horizons.

Overall, Table 3.11 indicates that the return association of the GAAP-based ICC component is strongest precisely when forecast definitions diverge most, with the amplification concentrated at short horizons.

Table 3.11: High earnings-divergence states: implied marginal effects

	Horizon (annualized)		
	12m	24m	36m
LowDiv			
ICC^{GAAP}	0.720***	0.227**	0.002
	(5.28)	(2.12)	(0.02)
ICC^{Diff}	0.898***	0.494**	0.288
	(2.75)	(2.18)	(1.44)
HighDiv			
ICC^{GAAP}	1.347***	0.319**	-0.025
	(3.93)	(2.29)	(-0.25)
ICC^{Diff}	0.944**	0.122	-0.039
	(2.36)	(0.57)	(-0.25)
Observations	95,254	94,752	94,251
Adj. R^2	0.176	0.186	0.178

This table reports implied marginal effects from predictive regressions of annualized future returns on a GAAP-based composite ICC, ICC^{GAAP}, and the Street–GAAP ICC differential, $ICC^{Diff}_{i,t} \equiv ICC^{Street}_{i,t} - ICC^{GAAP}_{i,t}$, allowing these effects to vary with forecast-definition divergence. HighDiv is an indicator equal to one for firm-months in the top quintile (within month) of average absolute forecast divergence, computed as the mean over horizons $h \in \{F1, F2, F5\}$ of $|(Frc_{Street,h} - Frc_{GAAP,h})/Price|$, where $Frc_{Street,h}$ and $Frc_{GAAP,h}$ denote Street and GAAP forecasts at horizon h. Forecasts are from I/B/E/S: Street forecasts correspond to I/B/E/S *EPS* fields, and GAAP forecasts correspond to I/B/E/S *GPS* (reported/GAAP) fields. The rows labeled *LowDiv* report the baseline marginal effects when HighDiv= 0 (main coefficients). The rows labeled *HighDiv* report marginal effects when HighDiv= 1, computed as the relevant linear combinations of main and interaction terms; statistical significance for these HighDiv effects is based on Wald tests. All specifications include controls for size (log market capitalization) and book-to-market, as well as industry and time fixed effects. Standard errors are heteroskedasticity-robust and clustered at the firm level. ***, **, and * denote significance at the 1%, 5%, and 10% levels.

3.5.5 Cross-sectional heterogeneity

Tables 3.12 and 3.13 examine whether the return associations documented above vary across settings commonly linked to informational frictions.

Table 3.12 reports return regressions estimated separately by size terciles (formed within month). The coefficient on ICC^{GAAP} is largest among smaller firms within the S&P 500 sample (1.049) and declines monotonically with size, becoming statistically insignificant for the largest tercile (0.238). The coefficient on the Street–GAAP ICC differential remains positive across terciles and is statistically meaningful even among larger firms (0.656). This pattern is consistent with GAAP-consistent expectations being relatively more informative when informational frictions are higher, while the definition-driven component embedded in Street inputs remains economically relevant even for widely followed firms.

Table 3.12: GAAP ICC and the Street–GAAP ICC differential across firm size terciles

	(1) Small	(2) Medium	(3) Big
ICC^{GAAP}	1.049***	0.621**	0.238
	(3.57)	(2.54)	(1.02)
ICC^{Diff}	0.760*	0.625*	0.656**
	(1.85)	(1.85)	(1.97)
Observations	28,488	31,686	35,080
Adj. R^2	0.227	0.168	0.183

Size terciles are formed within month by market capitalization.is the firm's subsequent 12-month buy-and-hold annual total return, measured from $t+1$ to $t+12$ relative to the month-t ICC estimate. ICC^{GAAP} is the composite implied cost of capital computed as the simple average of five valuation-model ICC estimates (GG, GLS, CT, MPEG, and OJ) using GAAP analyst forecast inputs. The Street–GAAP ICC differential is defined as $ICC^{Diff}_{i,t} \equiv ICC^{Street}_{i,t} - ICC^{GAAP}_{i,t}$ for the same firm-month, where ICC^{Street} uses I/B/E/S Street (EPS) forecasts and ICC^{GAAP} uses I/B/E/S reported (GPS/GAAP) forecasts. All specifications include industry and time fixed effects and controls. Standard errors are clustered by firm. Coefficients are reported in levels. ***, **, and * denote significance at the 1%, 5%, and 10% levels.

Table 3.13 examines industry heterogeneity using a partition by intangible intensity. This partition is motivated by evidence that intangible-intensive business models are harder to capture with standardized accounting measures and are associated with a more challenging information environment, so forecast-definition differences may be more consequential in these industries (e.g., Lev and Zarowin, 1999; Eisfeldt and Papanikolaou, 2013). The table splits the sample into a high-intangible group (Information Technology, Communication Services, and Health Care) and a low-intangible group (Materials, Industrials, Utilities, and Real Estate) and estimates the same return regression separately within each group.[2]

In high-intangible sectors, both components are positively associated with subsequent returns: the coefficient on ICC^{GAAP} is 0.805 (statistically significant) and the coefficient

[2]Within-group regressions include time fixed effects; industry fixed effects are not included because the sample is restricted to a small set of industries by construction.

on the Street–GAAP ICC differential is 0.918 (statistically significant). In low-intangible sectors, neither component is statistically distinguishable from zero ($ICC^{GAAP} = 0.325$; Street–GAAP ICC differential $= -0.064$). In this partition, the return associations linked to forecast-definition content are therefore concentrated in intangible-intensive sectors and are weak in more tangible, asset-based sectors.

Table 3.13: Industry heterogeneity: GAAP ICC and the Street–GAAP ICC differential

	High Intangible	Low Intangible
ICC^{GAAP}	0.805***	0.325
	(2.70)	(1.63)
ICC^{Diff}	0.918*	-0.064
	(1.77)	(-0.17)
Observations	25,846	29,111
Adj. R^2	0.156	0.197

The dependent variable is the firm's subsequent 12-month buy-and-hold annual total return, measured from $t+1$ to $t+12$ relative to the month-t ICC estimate. ICC^{GAAP} is the composite implied cost of capital computed as the simple average of five valuation-model ICC estimates (GG, GLS, CT, MPEG, and OJ) using GAAP analyst forecast inputs. The Street–GAAP ICC differential is defined as $ICC^{Diff}_{i,t} \equiv ICC^{Street}_{i,t} - ICC^{GAAP}_{i,t}$ for the same firm-month, where ICC^{Street} uses I/B/E/S Street (EPS) forecasts and ICC^{GAAP} uses I/B/E/S reported (GPS) forecasts. Columns report separate regressions estimated within two broad industry groups: High Intangible sectors comprise Information Technology, Communication Services, and Health Care; Low Intangible sectors comprise Materials, Industrials, Utilities, and Real Estate. Within each group, specifications include time fixed effects and the standard controls (log market capitalization and book-to-market, measured at t). Standard errors are heteroskedasticity-robust and clustered at the firm level. ***, **, and * denote significance at the 1%, 5%, and 10% levels.

3.5.6 Model-level evidence

The composite results can mask heterogeneity across ICC implementations. I therefore replicate the main return regressions separately for each valuation model (CT, GLS, MPEG, OJ, and GG), as well as for the composite. Table 3.14 reports model-by-model "horse-race" regressions at the 12-month horizon, entering ICC^{Street} and ICC^{GAAP} either one at a time or jointly. Table 3.15 re-parameterizes the same model-by-model analysis in terms of the GAAP-based ICC and the Street–GAAP ICC differential and traces the corresponding return associations across 12, 24, and 36 months. Finally, Tables 3.16 and 3.17 report model-level heterogeneity in the 12-month regressions across (i) firm-size terciles and (ii) high- versus low-intangible industries.

Table 3.14 shows that, when entered separately, both forecast definitions are positively associated with subsequent returns across all five implementations. When both definitions are included jointly, incremental content depends on valuation structure. In the abnormal-earnings-growth models (MPEG and OJ), the GAAP-based ICC retains incremental return-relevant variation when paired with the Street-based ICC. In contrast, incremental GAAP content is weaker in CT and GG, while GLS exhibits particularly high alignment between the two definitions and no significant incremental separation in the joint regression.

Table 3.14: Return regressions by ICC model: Street versus GAAP

	Street only	GAAP only	Joint		
Model	β(Street)	β(GAAP)	β(Street)	β(GAAP)	N
CT	0.878***	0.713***	0.917***	-0.040	93,055
	(4.76)	(4.39)	(3.41)	(-0.21)	
GLS	0.550***	0.562***	0.293	0.278	92,894
	(4.03)	(4.07)	(0.64)	(0.64)	
MPEG	0.569***	0.457***	0.357**	0.224**	84,012
	(5.15)	(5.13)	(2.39)	(1.98)	
OJ	0.575***	0.451***	0.382**	0.203*	83,935
	(5.18)	(5.03)	(2.53)	(1.77)	
GG	0.541***	0.513***	0.485*	0.043	91,361
	(3.37)	(3.57)	(1.72)	(0.18)	
Composite	0.928***	0.800***	0.719***	0.232	95,254
	(5.29)	(5.27)	(2.63)	(1.03)	

The dependent variable is the firm's subsequent 12-month buy-and-hold annual total return, measured from $t+1$ to $t+12$ relative to the month-t ICC estimate. Each row reports predictive regressions using ICCs computed from the indicated valuation model (GG, GLS, CT, MPEG, and OJ) and, in the last row, the composite ICC defined as the simple average across the five model-specific ICC estimates. "Street" and "GAAP" denote ICCs constructed using Street (analyst-adjusted) earnings forecasts versus reported (GAAP) earnings forecasts, respectively. Forecast inputs are from I/B/E/S: Street forecasts correspond to I/B/E/S *EPS* forecast fields; GAAP forecasts correspond to I/B/E/S *GPS* (reported/GAAP) forecast fields. Columns (1) and (2) include only the corresponding ICC measure, while the "Joint" specification includes both Street and GAAP ICCs simultaneously to assess incremental return-relevance. All specifications include controls for size (log market capitalization) and book-to-market, as well as industry and time fixed effects. Standard errors are heteroskedasticity-robust and clustered at the firm level. ***, **, and * denote significance at the 1%, 5%, and 10% levels.

Table 3.15 complements this evidence by expressing each Street-based ICC as a GAAP-based level plus a Street–GAAP ICC differential and estimating the return regressions at longer horizons. The evidence is concentrated at the 12-month horizon: for most models, both the GAAP component and the definition-driven differential are most reliably related to one-year-ahead returns, with weaker and less systematic associations at 24 months and generally small effects by 36 months.

Tables 3.16 and 3.17 show that these model-level patterns also vary across informational environments. Table 3.16 indicates that several implementations exhibit stronger GAAP-based coefficients among smaller firms, consistent with GAAP-consistent expectations being relatively more informative when valuation is harder and informational frictions are greater, while the Street–GAAP ICC differential remains economically rele-

Table 3.15: Return regressions by ICC model: GAAP ICC and the Street–GAAP ICC differential across horizons

	12m		24m		36m	
Model	ICC^{GAAP}	ICC^{Diff}	ICC^{GAAP}	ICC^{Diff}	ICC^{GAAP}	ICC^{Diff}
CT	0.877***	0.917***	0.328***	0.360**	0.055	0.153
	(4.52)	(3.41)	(2.98)	(2.03)	(0.58)	(1.04)
GLS	0.571***	0.293	0.158*	-0.416	-0.032	-0.472*
	(4.00)	(0.64)	(1.69)	(-1.23)	(-0.37)	(-1.69)
MPEG	0.581***	0.357**	0.083	0.006	-0.066	-0.060
	(5.12)	(2.39)	(1.24)	(0.06)	(-1.27)	(-0.83)
OJ	0.585***	0.382**	0.083	0.013	-0.065	-0.055
	(5.13)	(2.53)	(1.24)	(0.13)	(-1.24)	(-0.75)
GG	0.528***	0.485*	0.209**	0.147	0.032	-0.062
	(3.39)	(1.72)	(2.02)	(0.83)	(0.36)	(-0.41)
Composite	0.951***	0.719***	0.226**	0.154	-0.044	0.027
	(5.24)	(2.63)	(2.27)	(0.87)	(-0.54)	(0.20)

The dependent variable is the firm's subsequent buy-and-hold annual total return over the stated horizon, reported in annualized terms (12m: 12-month return; 24m and 36m: multi-year cumulative returns annualized to a per-year rate). Regressors are measured at month t, and returns are measured from $t+1$ onward. Each row corresponds to a valuation-model ICC (CT, GLS, MPEG, OJ, GG) and the composite ICC (simple average across the five model-specific ICC estimates). For each model and horizon, the regression is $r_{i,t\to t+h} = \alpha + \beta_1\, ICC^{GAAP}_{i,t} + \beta_2\, ICC^{Diff}_{i,t} + X_{i,t}\gamma + \mathrm{FE} + \varepsilon_{i,t+h}$, where $ICC^{Diff}_{i,t} \equiv ICC^{Street}_{i,t} - ICC^{GAAP}_{i,t}$ is the model-specific Street–GAAP ICC differential for the same firm-month. ICC^{Street} is computed using I/B/E/S Street (EPS) forecast fields, while ICC^{GAAP} is computed using I/B/E/S reported (GPS/GAAP) forecast fields. Controls $X_{i,t}$ include log market capitalization and book-to-market. Fixed effects are industry and time fixed effects. Standard errors are heteroskedasticity-robust and clustered at the firm level. ***, **, and * denote significance at the 1%, 5%, and 10% levels.

vant across a broader set of environments but with model-specific strength.

Table 3.17 provides a complementary split by intangible intensity. In high-intangible sectors, the return associations are concentrated in the growth-sensitive implementations: MPEG and OJ yield positive and statistically significant coefficients for both the GAAP-based ICC and the Street–GAAP ICC differential, while CT, GLS, and GG exhibit weaker or statistically indistinct relations. In low-intangible sectors, the pattern shifts. CT displays a strong positive GAAP coefficient and a positive Street–GAAP ICC differential, whereas the corresponding composite coefficients are small and statistically weak. This contrast illustrates that composite aggregation can attenuate model-specific relations when effects differ across implementations and partially offset in the average.

A valuation-based interpretation is consistent with these patterns. In intangible-intensive sectors, a larger share of value is tied to expectations about longer-run profitability and growth, so implementations that load more directly on implied growth (MPEG and OJ) are more responsive to definition-driven differences in forecast paths. In more tangible, accounting-anchored sectors, residual-income implementations that place

Table 3.16: Model-level size heterogeneity

	Small		Medium		Big	
Model	ICC^{GAAP}	ICC^{Diff}	ICC^{GAAP}	ICC^{Diff}	ICC^{GAAP}	ICC^{Diff}
CT	1.048***	0.859*	0.671***	0.762**	0.209	0.870***
GLS	0.585**	0.726	0.652***	-0.511	-0.056	-0.156
MPEG	0.635***	0.340	0.232	0.106	0.268*	0.377*
OJ	0.626***	0.349	0.244	0.142	0.276*	0.424**
GG	0.597**	1.254***	0.773***	0.563	-0.146	-0.860*
Composite	1.049***	0.760*	0.621**	0.625*	0.238	0.656**

Size terciles are formed within month based on market capitalization. The dependent variable is the firm's subsequent 12-month buy-and-hold annual total return, measured from $t+1$ to $t+12$ relative to the month-t ICC estimate. For each valuation model (CT, GLS, MPEG, OJ, GG) and for the composite ICC (simple average across the five model-specific ICC estimates), each tercile column reports coefficients from regressions of the form $r_{i,t\to t+12} = \alpha + \beta_1\, ICC^{GAAP}_{i,t} + \beta_2\, ICC^{Diff}_{i,t} + X_{i,t}\gamma + \text{FE} + \varepsilon_{i,t+12}$, estimated within the tercile. The model-specific Street–GAAP ICC differential is defined as $ICC^{Diff}_{i,t} \equiv ICC^{Street}_{i,t} - ICC^{GAAP}_{i,t}$ for the same firm-month. ICC^{Street} is computed using I/B/E/S Street (EPS) forecast fields, while ICC^{GAAP} is computed using I/B/E/S reported (GPS/GAAP) forecast fields. Controls $X_{i,t}$ include log market capitalization and book-to-market (measured at t). Fixed effects are industry and time fixed effects. Standard errors are heteroskedasticity-robust and clustered at the firm level. ***, **, and * denote significance at the 1%, 5%, and 10% levels.

greater weight on book values and accounting-consistent profitability (CT in particular) can deliver sharper return-relevant signal, and the role of forecast-definition differences can vary across models.

Table 3.17: Model-level industry heterogeneity

	High Intangible		Low Intangible	
Model	ICC^{GAAP}	ICC^{Diff}	ICC^{GAAP}	ICC^{Diff}
CT	0.295	0.015	0.668***	1.036**
GLS	0.093	-0.094	0.428***	-1.760
MPEG	0.634***	0.742**	0.138	-0.166
OJ	0.620***	0.749**	0.154	-0.113
GG	0.016	0.027	0.443*	0.545
Composite	0.805***	0.918*	0.325	-0.064

The dependent variable is the firm's subsequent 12-month buy-and-hold annual total return, measured from $t+1$ to $t+12$ relative to the month-t ICC estimate. Columns report separate regressions estimated within two broad industry groups: High Intangible sectors comprise Information Technology (GICS 45), Communication Services (50), and Health Care (35); Low Intangible sectors comprise Materials (15), Industrials (20), Utilities (55), and Real Estate (60). For each valuation model (CT, GLS, MPEG, OJ, GG) and for the composite ICC (simple average across the five model-specific ICC estimates), coefficients are from regressions of the form $r_{i,t\to t+12} = \alpha + \beta_1\, ICC^{GAAP}_{i,t} + \beta_2\, ICC^{Diff}_{i,t} + X_{i,t}\gamma + \text{FE} + \varepsilon_{i,t+12}$. The model-specific Street–GAAP ICC differential is defined as $ICC^{Diff}_{i,t} \equiv ICC^{Street}_{i,t} - ICC^{GAAP}_{i,t}$ for the same firm-month. ICC^{Street} is computed using I/B/E/S Street (EPS) forecast fields, while ICC^{GAAP} is computed using I/B/E/S reported (GPS/GAAP) forecast fields. Controls $X_{i,t}$ include log market capitalization and book-to-market (measured at t). Within each industry group, specifications include year fixed effects. Standard errors are heteroskedasticity-robust and clustered at the firm level. ***, **, and * denote significance at the 1%, 5%, and 10% levels.

3.6 Conclusion

This paper shows that the earnings forecast definition used to construct implied costs of capital is a first-order measurement decision. Using Street versus GAAP analyst forecasts changes ICC levels and ranks. Because many ICC applications rely on ranks, tails, and cutoffs rather than on levels, forecast definition can shift expected return inference in designs that sort, condition, or interact on ICCs.

The evidence clarifies why definition choice matters. Street and GAAP forecasts differ systematically in both near term levels and the growth profile implied across horizons, and standard implied cost of capital implementations translate these differences into implied discount rates in model-specific ways. As a result, even when Street- and GAAP-based ICCs are highly correlated on average, switching the forecast definition can materially reassign firms in the tails of the ICC distribution. A decomposition that separates the GAAP-based ICC from the Street–GAAP ICC differential sharpens interpretation and helps distinguish accounting consistent expectations from the incremental content associated with exclusions. In return association tests, the GAAP-based component is most informative when definitions diverge most and in settings where informational frictions are plausibly greater, including smaller firms and intangible intensive industries. Measurement error variance diagnostics indicate broadly similar aggregate reliability for Street- and GAAP-based constructions, so the key issue is not whether one definition is admissible and the other is not, but whether the two definitions capture different economic content in different settings.

For empirical work, the implication is straightforward. Studies should state the forecast definition explicitly, treat it as part of the measurement design, and report sensitivity to the alternative definition when feasible, especially when inference comes from portfolio sorts, tail indicators, or threshold-based conditioning. When the objective is to isolate definition related variation, the Street–GAAP ICC differential provides a structured object that maps forecast definition into implied discount rate variation while holding prices and non-forecast inputs fixed.

Two directions for future research follow. The first is external validity. The analysis focuses on large-cap U.S. firms with relatively rich forecast coverage. An important next step is to examine broader universes where coverage is thinner and reporting and information frictions are more heterogeneous, and to test whether definition effects and tail reshuffling are amplified in those settings. The second is incentives. If Street-based inputs are treated as a default in valuation and expected return applications, firms may have incentives to influence the Street relative to GAAP mapping. This can be tested by linking the definition related implied cost of capital component to observable non-GAAP reporting choices, such as the frequency and persistence of adjustments, their salience in earnings communications, and changes in the classification of excluded items.

Chapter 4

Beyond ESG Scores: Risk and Return Insights from SFDR Metrics‡

4.1 Introduction

Environmental, social, and governance (ESG) considerations have become increasingly central to investment decision-making and financial regulation. Institutional investors, asset managers, and regulators alike seek to align capital allocation with sustainability objectives while preserving financial performance. In response, sustainable finance has evolved into a broad set of investment approaches that incorporate ESG-related information into portfolio construction, ranging from exclusionary screening and best-in-class selection to full ESG integration and impact-oriented strategies (CFA Institute et al. (2023)).

Beyond their normative appeal, ESG criteria have attracted substantial attention due to their potential financial relevance. A large empirical literature examines whether firms with superior ESG performance exhibit differences in risk, return, and resilience compared to conventional firms. While many studies document a negative association between ESG engagement and firm risk, evidence on whether ESG integration improves risk-adjusted returns remains mixed. Meta-analyses suggest that ESG integration is, on average, not negative related to financial performance, but portfolio-level studies often find weaker or heterogeneous results, reflecting differences in investment strategies, implementation costs, data providers, and methodological choices (e.g., Friede et al.

‡This paper is based on Conrads and Zährl (2026). We thank Mario Hendriock and Dieter Hess for helpful comments and suggestions.

(2015); Gillan et al. (2021); Pedersen et al. (2021)).

A prominent strand of the literature highlights firm risk as a key channel through which ESG characteristics may affect financial performance. Studies show that firms with stronger ESG or CSR profiles tend to face lower financing costs and lower downside or firm-specific risk (El Ghoul et al. (2011); Hoepner et al. (2024)). These effects appear particularly pronounced during periods of market stress, when firms with stronger environmental and social profiles have been shown to exhibit greater resilience (Albuquerque et al. (2020)). At the same time, reduced risk does not necessarily translate into higher Sharpe ratios, because expected returns may also adjust when investors place value on sustainability characteristics themselves (Pedersen et al. (2021)).

A central challenge in assessing the financial implications of ESG investing lies in the measurement of ESG performance itself. ESG ratings provided by commercial data vendors often diverge substantially across providers, reflecting differences in indicator selection, weighting schemes, aggregation methods, and normative perspectives (Berg et al. (2022)). Conventional ESG scores typically combine positive, neutral, and negative firm activities into a single composite measure, which may obscure firms' actual societal and environmental externalities and introduce considerable noise into empirical analyses. Recent evidence suggests that methodological design choices, particularly aggregation and normalization procedures, play a major role in driving ESG score variation, sometimes outweighing differences in underlying firm disclosures (e.g., Berg et al. (2022); Benuzzi et al. (2025); Guerrero and Viteri (2025)).

Against this background, this paper adopts an adverse-impact-oriented approach to ESG measurement. Rather than relying on broad composite ESG scores, we focus exclusively on negative externalities as captured by the Principal Adverse Impact (PAI) indicators defined under the European Union's Sustainable Finance Disclosure Regulation (SFDR). These indicators provide a standardized and regulation-based framework that directly measures the harmful environmental and social effects of corporate activities. By concentrating on adverse impacts, we aim to isolate the dimensions of ESG performance that are most closely linked to regulatory, reputational, and financial risk.

Methodologically, we further address an important but underexplored source of ESG rating divergence: aggregation and normalization. Building on recent evidence that percentile-rank ESG scores can produce distorted and "sticky" outcomes (Benuzzi et al. (2025)), we compare conventional percentage-rank metrics with performance-rank (PR) metrics, which more directly reflect firms' relative sustainability performance. Using data from the London Stock Exchange Group (LSEG), we construct an SFDR-based score from PAI indicators and evaluate portfolio strategies based on both percentile and performance-rank methodologies. We benchmark these portfolios against strategies formed on LSEG's standard ESG and ESG Combined Score (ESGC) scores. Our results show that the SFDR PR outperforms all other score metrics, delivering return spreads of up to 0.77 percentage points and superior Sharpe ratios. More generally, the performance-rank method yields return spreads of up to 0.34 percentage points higher than its percentage-rank counterpart and increases Sharpe ratios by up to 0.19 under a best-in-class investment approach.

This paper makes two main contributions. First, we translate SFDR PAI information into an investable firm-level metric that offers a targeted alternative to broad composite ESG scores. Second, we examine how methodological design choices in sustainability measurement shape portfolio results across both conventional ESG scores and PAI-based metrics. We show that performance-based rankings generate stronger return spreads than percentile-based scores, underscoring the importance of construction choices when sustainability information is used in portfolio formation.

The remainder of the paper is structured as follows. We begin in Section 4.2 by reviewing the related literature and positioning our study within the existing research. Section 4.3 then introduces the data, variables, and the construction of the sustainability metrics employed in the analysis. Building on this framework, Section 4.4 examines the relationship between sustainability metrics and financial performance. Finally, Section 4.5 concludes.

4.2 Related literature

A substantial body of research examines the relationship between corporate social responsibility (CSR), ESG performance, and firm risk. While CSR and ESG differ conceptually, with CSR emphasizing corporate policies and stakeholder engagement and ESG focusing on measurable sustainability outcomes, we follow the literature in treating them as closely related constructs. Early work by Godfrey et al. (2009) argues that CSR generates goodwill among stakeholders, providing insurance-like protection against adverse events. Consistent with this view, Jo and Na (2012) document an inverse relationship between CSR engagement and firm risk, particularly in controversial industries, while Oikonomou et al. (2012) show that firms with strong ESG characteristics adopt more advanced risk management practices. Research on systematic risk further indicates that firms with higher environmental efficiency and transparency are less vulnerable to market-wide shocks and benefit from lower information asymmetry (e.g., Eccles et al. (2014); El Ghoul et al. (2011)). More recent studies confirm that high ESG ratings are associated with lower downside risk, fewer idiosyncratic incidents, and reduced earnings volatility (e.g., Giese et al. (2019); Hoepner et al. (2024)). While this literature establishes that ESG engagement can mitigate risk, it provides limited guidance on which ESG dimensions drive these effects. Our study addresses this gap by focusing explicitly on adverse-impact indicators.

A related strand investigates whether ESG performance translates into superior financial outcomes. Meta-analyses by Friede et al. (2015) and Whelan et al. (2021) find that most studies report a non-negative association between ESG integration and financial performance, although results vary across regions, time periods, and methodologies. Portfolio-level analyses often produce weaker or mixed findings, inconsistent ESG strategies, and measurement challenges. Theoretical frameworks, such as the efficient ESG frontier proposed by Pedersen et al. (2021), illustrate how ESG constraints affect optimal portfolios and Sharpe ratios. Empirical evidence by Prol and Kim (2022) suggests that while high-ESG portfolios exhibit lower volatility, this reduction is often accompanied by lower returns, yielding ambiguous effects on risk-adjusted performance. Although prior literature emphasizes the importance of ESG measurement choices, the portfolio implications of specific design features, such as aggregation or ranking methodology, remain less well understood. Our paper contributes by systematically comparing percentile- and performance-based ESG metrics in a portfolio context.

Several studies examine ESG performance during periods of financial distress, when risk mitigation is particularly salient. Lins et al. (2017) show that firms with high ESG ratings accumulated social capital prior to the Global Financial Crisis, enabling better access to financing and higher profitability during the downturn. Additional evidence indicates that high-CSR firms raised debt more easily and at lower cost during crises (e.g., Cornett et al. (2016); Amiraslani et al. (2023)). Evidence from the COVID-19 pandemic is more mixed: some studies report greater resilience and lower financial risk among high-ESG firms (e.g., Albuquerque et al. (2020); Broadstock et al. (2021); Engelhardt et al. (2021)), whereas others find limited or no abnormal return effects (e.g., Bae et al. (2021); Demers et al. (2021)). These mixed results highlight that investors may distinguish between substantive ESG engagement and symbolic or opportunistic behavior, suggesting the need for precise measures of ESG performance, such as those focused on adverse impacts.

Another important strand highlights divergence among ESG rating providers. Chatterji et al. (2016) attribute differences to definitions and measurement approaches, while Berg et al. (2022) report low correlations across major ESG vendors. Geographic focus, legal traditions, and normative perspectives further contribute to rating dispersion (See Gibson Brandon et al. (2021)). Beyond indicator selection and weighting, aggregation and normalization procedures can distort ESG rankings. Benuzzi et al. (2025) show that percentile-ranking methods inflate scores for top firms and generate path-dependent outcomes, with a large share of score variation driven by methodological design rather than underlying firm behavior. While this literature shows that methodological choices materially affect ESG scores and rankings, much less is known about how aggregation and normalization choices translate into portfolio performance. Addressing this link is a central focus of our analysis.

Finally, concerns about the limited comparability of broad ESG ratings have also motivated regulatory efforts to standardize sustainability-related disclosure. In the European Union, the Sustainable Finance Disclosure Regulation (SFDR) and its Regulatory Technical Standards define a standardized set of Principal Adverse Impact (PAI) indicators and disclosure templates intended to improve the transparency and comparability of sustainability-related information (European Parliament and the Council of the European Union (2019); European Commission (2022); Busch (2023)). Unlike conventional ESG scores, which typically combine positive and negative attributes using provider-specific methodologies, PAI indicators adopt a common adverse-impact perspective and therefore offer a more transparent basis for assessing sustainability-related downside exposures. Existing research, however, has focused mainly on SFDR at the fund and disclosure level. Prior studies show that SFDR classifications affect fund sustainability characteristics and investor flows (Becker et al. (2022)), that Article 9 funds remain heterogeneous in their actual impact orientation (Scheitza and Busch (2024)), and that the regulation can mitigate greenwashing among some sustainable funds (Abouarab et al. (2025)). Much less is known about how regulator-defined PAI indicators can be combined into firm-level scores, how alternative aggregation and ranking procedures affect those scores, and how such design choices translate into portfolio performance. Our study addresses this gap by constructing firm-level PAI-based measures and evaluating their portfolio implications under alternative scoring approaches.

4.3 Data and methodology

The following section provides the methodological foundation for the empirical examination of the performance of SFDR- and ESG-based metrics. Section 4.3.1 provides a detailed description of the dataset and the variables used in the analysis. Section 4.3.2 explains the procedures employed to calculate the sustainability metrics, while Section 4.3.3 presents descriptive statistics summarizing the distribution and variation of the SFDR and ESG metrics.

4.3.1 Data

The dataset originates from the LSEG database and contains year-end data for all companies that have ESG coverage available for the period from 2002 to 2023. The dataset combines financial and sustainability data at the company level, including ESG and ESGC scores, indicator scores, and a proprietary, equally weighted ESG score. This score is created using a uniform weighting methodology of all 186 ESG data points. Additionally, we use indicators relevant to the PAI statement to construct an SFDR score that captures firm performance in line with the SFDR. Financial performance is measured using total return. Overall, the data set covers a total of 97,054 firm-year observations.

4.3.2 Score calculation

The following section explains how the sustainability metrics are formed and calculated.

The ESG and ESGC scores are taken directly from LSEG. Our new SFDR score is calculated using the same method as the LSEG scores. The score is based on the PAI statement variables listed in the Appendix C.1.

First, a relative score is calculated for each data point based on firm performance within the same year and industry. For each observation, the score is defined as the proportion of firms with lower values (worse_count) plus half the proportion of firms with equal values (same_count), divided by the total number of firms (n):

$$\text{Score} = \frac{\text{worse\_count} + 0.5 \times \text{same\_count}}{n}. \tag{4.1}$$

This approach ensures that polarity is correctly aligned: higher values are assigned to better performance when the indicator has positive polarity and to worse performance when the indicator has negative polarity. The resulting per-indicator scores were then summed across all data points to obtain an aggregated firm-level score for each year and industry. In a second step, this procedure was reapplied to the aggregated scores to generate a final normalized score (SFDR Score) consistent with the LSEG ESG scoring methodology.

In light of the criticism by Benuzzi et al. (2025) regarding the aggregation of ESG scores and their normalization, we also calculate a SFDR score based on performance rank values (SFDR PR). We evaluate the indicator values again for each year and industry. We normalize the respective data points to a value between 0 and 100, resulting in the PR:

$$\text{PR} = \frac{x - \min(x)}{\max(x) - \min(x)} \times 100. \tag{4.2}$$

In addition, we calculate an equally weighted ESG Score (ESG EW). The score is based on all 186 ESG data points which are ranked through LSEG's percentile rank method and aggregated through an equal weighting. Moreover, we also calculate a ESG EW PR metric which follows the calculation of the SFDR PR. Thus, our analysis includes the following six sustainability metrics: ESG, ESGC, SFDR, SFDR PR, ESG EW, and ESG EW PR.

4.3.3 Descriptive statistics

A first look at the data is shown in Table 4.1, which presents the descriptive statistics for the firm-level ESG and SFDR metrics over the period examined from 2002 to 2023. The metrics include the raw ESG and ESGC scores, the SFDR score and its corresponding PR, as well as the proprietary equally weighted ESG score and its PR. For each metric, the table reports the mean, median, standard deviation, and key percentiles across the sample. These statistics allow for an understanding of both the central tendency and dispersion of the scores.

The mean ESG score is 43.62, with values ranging from 16.87 at the 10th percentile (P10) to 72.65 at the 90th percentile (P90), indicating a relatively wide range of performance across the sample. As expected, the ESGC score is slightly lower because it includes a discount factor for controversies, with a mean of 42.36 and a range from 16.79 (P10) to 69.79 (P90). In contrast, the SFDR and ESG EW scores are more evenly distributed, with mean values around 50, a P10 of approximately 10, and a P90 near 90. This reflects the double percentile ranking procedure and the equal weighting of individual indicators, which spreads companies across the full scale of 0 to 100. PR Scores show substantially lower values. For instance, the mean SFDR PR is 20.77, indicating that, on average, firms lag significantly behind top-ranking peers in sustainability performance. The distribution is also shifted left compared to the percentile scores, with P10 at 4.22 and P90 at 40.46, highlighting the gap between average and top-performing entities. Similarly, the ESG EW PR has a mean of 29.19, which is higher than the SFDR PR, yet still considerably lower than the ESG and ESG EW scores. This suggests that although many companies are in the middle of the distribution in terms of regular ESG scores, they are relatively far from the highest-performing peers in terms of absolute performance. In other words, the ESG and ESG EW scores reflect relative ranking, or how a company compares to its peers, while PR scores capture performance distance, or how much a company lags behind the top firm. In that sense, PR scores highlight meaningful performance gaps that would have been overlooked with standard ESG scores.

Table 4.1: Descriptive statistics

Score	Mean	Std. Dev.	P10	P25	P50	P75	P90
ESG Score	43.6179	20.6813	16.8710	26.9386	42.3024	59.5608	72.6502
ESGC Score	42.3614	19.6731	16.7964	26.7145	41.3411	57.2045	69.7946
SFDR Score	50.2107	28.8394	10.1211	25.1397	50.1721	75.1397	90.0809
SFDR PR	20.7685	13.8780	4.2254	7.8130	18.8980	31.9777	40.4595
ESG EW Score	50.2503	28.8134	10.1542	25.2780	50.2841	75.1389	90.0000
ESG EW PR	29.1863	11.5707	14.5269	20.5630	28.3519	37.9739	45.0546

This table reports descriptive statistics for firm-level sustainability metrics from LSEG for the sample period 2002–2023. ESG Score and ESGC Score are LSEG proprietary composite ratings. SFDR Score is constructed from LSEG variables that map into the EU Sustainable Finance Disclosure Regulation Principal Adverse Impact (PAI) indicators and related adverse-impact fields used in the PAI reporting framework; the score aggregates these adverse-impact inputs into a single firm-level measure. "PR" denotes the corresponding performance-rank version of a metric: instead of assigning each firm a percentile position within the cross-section, PR converts underlying indicator values into ranks that preserve magnitude differences in performance before aggregation, thereby reducing the compression inherent in percentile-based scaling. "ESG EW" denotes an equally weighted variant of the ESG metric constructed from the same underlying components. The table reports the mean, standard deviation, and selected percentiles (P10, P25, P50, P75, P90) computed over firm-year observations.

A correlation analysis provides additional insight into how these sustainability indicators relate to one another. As shown in Table 4.2 both Pearson and Spearman coefficients are uniformly positive and generally high, indicating a strong degree of coherence across the different metrics, with correlations typically exceeding 0.69. The strongest correlation is between the ESG and ESGC scores (ρ = 0.97), which suggests that controversy adjustments influence score magnitudes, while leaving the underlying firm ranking largely unchanged. Generally, the standard ESG score shows strong correlations with other ESG-based metrics (ESGC, ESG EW, ESG EW PR) and somewhat lower correlations with SFDR metrics (ρ = 0.78–0.79). Percentile-rank scores align closely across frameworks, with SFDR PR and ESG EW PR correlating above 0.90. This shows that, despite differences in focus, the relative ordering of firms is largely preserved.

Overall, these patterns reveal two key insights: the impact of the methodological variation on sustainability metrics and the substantive differences between ESG and SFDR frameworks. These insights are expected to influence subsequent analyses of sustainability performance.

Table 4.2: Correlation matrix

Score	ESG Score	ESGC Score	SFDR Score	SFDR PR	ESG EW Score	ESG EW PR
ESG Score	1.0000	0.9734	0.7796	0.7937	0.8879	0.8634
ESGC Score	0.9650	1.0000	0.7607	0.7783	0.8569	0.8357
SFDR Score	0.7750	0.7555	1.0000	0.7796	0.8376	0.6999
SFDR PR	0.7861	0.7678	0.7709	1.0000	0.7156	0.9171
ESG EW Score	0.8861	0.8537	0.8377	0.7113	1.0000	0.8089
ESG EW PR	0.8648	0.8265	0.6940	0.9028	0.8094	1.0000

This table reports pairwise correlations among six firm-level sustainability metrics from LSEG over 2002–2023. Entries in the lower triangle are Pearson correlations based on metric levels, and entries in the upper triangle are Spearman rank correlations based on within-year firm rankings. ESG Score and ESGC Score are LSEG composite ratings. SFDR Score aggregates adverse-impact variables mapped to the EU SFDR Principal Adverse Impact (PAI) framework into a single firm-level measure. "PR" denotes the performance-rank version of a metric constructed from ranks of the underlying indicators prior to aggregation, which preserves performance gaps more directly than percentile scoring. "ESG EW" denotes an equally weighted variant of the ESG metric. Correlations are computed using firm-year observations.

4.4 Financial performance

In the following section, the main part of this paper, we evaluate the relationship between sustainability and financial performance. In Section 4.4.1, we provide an overview of returns and sustainability metrics across deciles, including the distribution of firms across deciles and the alignment of different sustainability metrics. This allows us to assess whether higher sustainability performance translates into observable differences in financial results. Then, in Section 4.4.2, we analyze investment strategies based on these sustainability metrics. Firms are sorted into portfolios based on the respective metrics, and we examine both the returns and risk-adjusted performance of these strategies to evaluate their practical relevance for investors and the extent to which sustainability can inform portfolio construction.

4.4.1 Returns across sustainability deciles

To gain an initial understanding of the relationship between sustainability and financial performance, firms are divided into deciles based on each sustainability metric. Table 4.3 reports the mean sustainability scores by decile, showing a clear progression from Decile 1 (D1) at the bottom to Decile 10 (D10) at the top. ESG scores increase from approximately 12 in D1 to 83 in D10, while SFDR scores span a wider range, from 5 to 95, reflecting the percentile-based normalization. PR scores are generally lower, highlighting the gap between average firms and top performers; in D10, the mean SFDR PR is 65, and ESG EW PR reaches 74.

Table 4.3: Mean score per decile

Score	1	2	3	4	5	6	7	8	9	10
ESG Score	12.1825	21.5983	28.3416	34.6182	40.9864	47.6823	54.6267	62.0257	70.5359	82.7815
ESGC Score	12.2387	21.6473	28.3386	34.4753	40.5643	46.7369	53.1158	60.0050	68.1560	80.4201
SFDR Score	5.0911	15.0975	25.1808	35.2752	45.3254	55.2891	65.2786	75.2236	85.1543	95.0579
SFDR PR	5.1699	9.9796	14.6265	20.4150	26.9932	33.8897	40.5487	46.9732	53.9722	65.3375
ESG EW Score	5.0528	15.1845	25.3019	35.3909	45.4249	55.3779	65.3248	75.2326	85.1187	95.0126
ESG EW PR	19.1911	28.0489	32.8425	37.1013	41.5793	46.5194	51.6360	57.0946	63.5858	74.2393

This table reports decile means for six firm-level sustainability measures used in the portfolio tests: the LSEG *ESG Score* and *ESG Combined Score (ESGC)*, an SFDR-based score constructed from Principal Adverse Impact (PAI) indicators (*SFDR Score*), and their alternative ranking/aggregation variants. Metrics labeled *PR* are based on the performance-rank transformation, whereas the non-PR versions use the baseline percentile-ranking approach; *EW* denotes an equal-weight aggregation across underlying indicators. Firms are sorted each period into deciles based on the respective metric, and the table reports the mean metric value within each decile. The sample spans 2002–2023.

To move from the distribution of sustainability metrics to their financial implications, we next examine mean annual returns for each decile, allowing us to assess whether sustainability rank, under either metric, is associated with systematic differences in firm performance.

Table 4.4 reports the mean annual returns for each decile across all sustainability metrics. Several patterns emerge. In general, returns increase from the lowest to the highest deciles, although the strength and monotonicity of this relationship vary across metrics. For the standard ESG score, mean annual returns rise from roughly 2.2% in D1 to about 9.8% in D10, suggesting a positive association between sustainability performance and realized returns. A similar pattern is visible for ESGC and SFDR scores, where higher-decile firms tend to earn higher average returns. For the PR-based metrics, the relationship is likewise upward-sloping but somewhat less smooth. The lowest SFDR PR decile stands out, exhibiting the weakest performance of all, indicating that firms classified as the "least sustainable" according to the SFDR PR metric also tend to generate the lowest returns. The middle and upper deciles show stronger and more consistent returns, with the top deciles again delivering the highest averages. These results indicate that across all metrics, firms classified as sustainability leaders tend to achieve higher realized returns relative to firms in the bottom deciles, while the weakest performers can be particularly penalized in terms of returns.

Table 4.4: Mean return per decile

Score	1	2	3	4	5	6	7	8	9	10
ESG Score	0.0218	0.0444	0.0613	0.0615	0.0756	0.0832	0.0800	0.0798	0.0877	0.0975
ESGC Score	0.0207	0.0446	0.0625	0.0633	0.0727	0.0829	0.0850	0.0771	0.0877	0.0965
SFDR Score	0.0327	0.0499	0.0692	0.0686	0.0742	0.0717	0.0801	0.0777	0.0774	0.0915
SFDR PR	0.0167	0.0332	0.0648	0.0732	0.0751	0.0801	0.0836	0.0836	0.0885	0.0940
ESG EW Score	0.0376	0.0505	0.0553	0.0726	0.0745	0.0720	0.0717	0.0757	0.0891	0.0939
ESG EW PR	0.0322	0.0222	0.0556	0.0777	0.0815	0.0727	0.0773	0.0840	0.0929	0.0968

This table reports mean annual returns across deciles formed on six firm-level sustainability measures: the LSEG *ESG Score* and *ESG Combined Score (ESGC)*, an SFDR-based score constructed from Principal Adverse Impact (PAI) indicators (*SFDR Score*), and their alternative ranking/aggregation variants. Metrics labeled *PR* use the performance-rank transformation, whereas the non-PR versions use the baseline percentile-ranking approach; *EW* denotes an equal-weight aggregation across underlying indicators. Firms are sorted each period into deciles based on the respective metric. The table reports the average annual return of firms within each decile over the sample period 2002–2023.

Since our broader goal is to understand how differences in sustainability measurement and scoring methodologies translate into portfolio outcomes, it is important to examine how firms are ordered across metrics. In particular, differences in firm rankings can lead to distinct portfolio compositions and potentially different return patterns. To illustrate this, Figure 4.1 presents a decile-by-decile heatmap of the cross-sectional overlap between standard ESG scores and SFDR PR scores. Each cell reports the percentage of the full sample that is simultaneously assigned to the corresponding ESG decile and SFDR PR decile. The figure reveals a clear concentration of firms along, but not exclusively on, the main diagonal: higher ESG deciles tend to contain a greater share of firms that also rank highly on the SFDR PR metric, indicating a positive association between the two measures of sustainability, with up to 49% of firms falling into the same decile across both measures. At the same time, the sizable share of observations away from the diagonal underscores that ESG and SFDR PR classifications diverge for many firms, capturing related but not identical dimensions of sustainability performance.

This classification pattern is central for the empirical strategy that follows. In the next section, we construct portfolio strategies based on these rankings and evaluate their performance using returns and Sharpe ratios. The imperfect alignment between ESG and SFDR PR rankings suggests that portfolios constructed using these metrics will differ in composition. As a result, any differences in returns or risk-adjusted performance may reflect not only firm sustainability but also the choice of measurement. Thus, the heatmap provides a visual guide for interpreting these differences and highlights how metric selection can influence portfolio outcomes.

4.4.2 Investment strategies

To assess how sustainability metrics translate into portfolio performance, we construct four core investment strategies that differ in the treatment of firms with high and low sustainability scores. Firms are classified using our six sustainability metrics and sorted into deciles, where D1 contains the firms with the lowest sustainability scores, while D10 contains firms with the highest scores. Throughout the analysis, we refer to D1 as "bad" firms and D10 as "good" firms, with the intermediate groups representing gradually improving sustainability profiles.

Strategy 1^b invests exclusively in firms classified as bad, while Strategy 1^g includes only good firms and therefore corresponds to a best-in-class approach, targeting only the

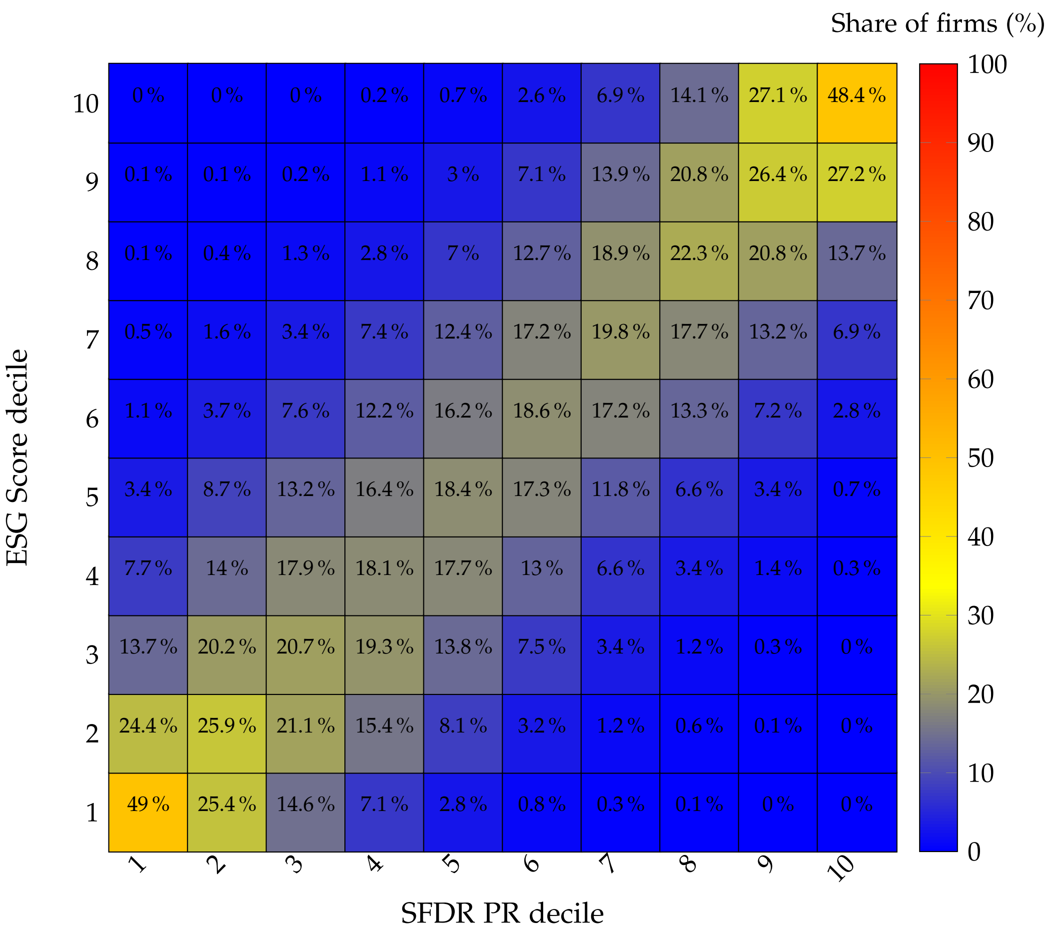


Figure 4.1: Overlap of firms across deciles

This figure reports the joint distribution of firms across deciles of *SFDR PR* (x-axis) and the *ESG Score* (y-axis) over the sample period 2002–2023. In each period, firms are independently sorted into deciles on each metric (Decile 1 = lowest values; Decile 10 = highest values). *SFDR PR* is constructed from the SFDR Principal Adverse Impact (PAI) indicators by (i) orienting each indicator so that higher values reflect *better* sustainability performance (i.e., lower adverse impact), (ii) computing firm ranks on each oriented indicator in the cross-section, and (iii) aggregating these indicator ranks into a composite performance-rank score (with equal weights across indicators). Each cell shows the percentage of firm–period observations that fall into the corresponding pair of deciles.

firms with the highest sustainability scores.

The second set of strategies implements exclusion rules. Strategy 2^b removes the "good" firms, i.e., D10, while maintaining equal investment in all other deciles. Conversely, Strategy 2^g excludes the "bad" firms, i.e., D1, and treats all the other deciles equally. Thus, it represents a negative-screening methodology.

Table 4.5 presents the annual mean returns associated with each metric and strategy. A consistent pattern emerges across all metrics: portfolios composed solely of "good" firms (Strategy 1^g) tend to achieve the highest average returns. For example, using the ESG Score, the "all-good" portfolio yields an annual mean return of 10.77%, compared with only 6.36% for the all-bad portfolio. Similar gaps appear for the ESGC and SFDR-based scores. Negative screening approaches deliver more moderate but still robust returns. Strategy 2^g, which excludes only the bad firms, performs slightly better than Strategy 2^b across most metrics, suggesting that removing the lowest-rated firms produces marginal improvements without requiring an extreme shift in portfolio composition.

Table 4.5: Annual mean returns

Score	Strategy 1^b	Strategy 1^g	Strategy 2^b	Strategy 2^g
ESG Score	0.0636	0.1077	0.0908	0.0957
ESGC Score	0.0630	0.1054	0.0911	0.0958
SFDR Score	0.0670	0.1076	0.0908	0.0954
SFDR PR	0.0530	0.1110	0.0905	0.0969
ESG EW Score	0.0680	0.1038	0.0913	0.0952
ESG EW PR	0.0599	0.1064	0.0910	0.0961

This table reports annual mean returns for portfolios formed on six firm-level sustainability metrics from LSEG over 2002–2023: ESG Score, ESGC Score, SFDR Score, SFDR PR, ESG EW Score, and ESG EW PR. ESG Score and ESGC Score are LSEG proprietary composite ratings; ESGC incorporates an additional adjustment for controversies relative to the standard ESG Score. SFDR Score is constructed by aggregating LSEG variables that map into the EU Sustainable Finance Disclosure Regulation Principal Adverse Impact (PAI) framework. "PR" denotes the performance-rank version of a metric, constructed by ranking firms on the underlying indicator inputs prior to aggregation. "ESG EW" denotes an equally weighted variant of the ESG metric based on the same underlying components. For each metric and year, firms are sorted into deciles, where Decile 1 contains the lowest-scoring firms ("bad") and Decile 10 the highest-scoring firms ("good"). Four portfolio strategies are considered: Strategy 1^b holds Decile 1 only, Strategy 1^g holds Decile 10 only, Strategy 2^b excludes Decile 10 and holds an equal-weighted portfolio of Deciles 1–9, and Strategy 2^g excludes Decile 1 and holds an equal-weighted portfolio of Deciles 2–10. Reported values are time-series averages of annual portfolio returns.

Table 4.6 reports the annual mean return spreads between the good and bad firms for the two main strategies. For Strategy 1, the spread is defined as the performance difference between the good-only portfolio (1^g) and the bad-only portfolio (1^b). For Strategy 2, the spread captures the difference between the negative screen that excludes bad firms (2^g) and the screen that excludes good firms (2^b). In both cases, a positive spread indicates that portfolios tilted toward firms with better sustainability characteristics outperform those tilted toward poorer performers. Across all six sustainability metrics, the spreads are positive, statistically significant at the 1% level, and economically meaningful.

Across all six sustainability metrics, the Strategy 1 spreads are economically large and consistently statistically significant. They range from 3.58% up to 5.80% per year, indicating that the good-only portfolios outperform the bad-only portfolios by a substantial margin. The largest spread occurs for the SFDR PR metric, where the annual difference reaches 5.80%, suggesting that the adverse-impact-based classification combined with the performance ranking generates the strongest performance separation between the two extremes of the sustainability distribution. By contrast, the smallest spread, though still considerable, appears for the ESG EW Score, amounting to 3.58%. The remaining metrics all produce spreads between roughly four and five percent, specifically 4.41% for the ESG Score, 4.24% for the ESGC Score, 4.06% for the SFDR Score, and 4.65% for the ESG EW PR score, demonstrating a remarkably consistent relationship between sustainability level and return outcomes when comparing the portfolio allocations.

A similar but quantitatively more muted pattern emerges for Strategy 2. Here, the return spreads lie between 0.40% and 0.64% per year, yet all are statistically significant at the 1% level. Although the economic magnitude is smaller than in the concentrated Strategy 1 comparison, the spreads nonetheless indicate that excluding the worst-scoring firms yields systematically higher returns than excluding the best-scoring firms. Once again, the SFDR PR metric produces the highest spread, at 0.64%, whereas the ESG EW score generates the lowest, at 0.40%. The remaining metrics fall within a tight range around one-half of one percentage point.

These results provide clear evidence that superior sustainability performance is associated with enhanced portfolio returns. Even when only partial exclusions are applied, higher-rated firms systematically deliver positive return premiums, indicating that integrating sustainability metrics into investment decisions can offer both financial and ethical advantages.

While Table 4.6 highlights the absolute performance spreads of different strategies, it does not directly reveal the incremental effects of using one sustainability metric over another. To capture these relative differences, Table 4.7 presents annual mean return spreads between pairs of sustainability metrics. This allows us to examine whether certain sustainability metrics consistently lead to higher returns under the same investment strategy. Importantly, we are interested in identifying whether these differences are statistically significant.

The results show that, for most pairwise comparisons, return spreads are small and statistically insignificant, suggesting that portfolio performance is largely robust across different sustainability metrics. For Strategy 1^g, the largest positive spread is observed between SFDR PR and ESG EW Score (0.72 p.p.), though it is not statistically significant. For Strategy 2^g, the most notable spread is also between SFDR PR and ESG EW Score (0.17 p.p.) and SFDR Score and SFDR PR (-0.15 p.p.), which reaches weak statistical significance. These results indicate that significant differences, when they occur, are primarily associated with the SFDR PR metric, which tends to outperform other scoring

Table 4.6: Annual mean return spreads

Score	Strategy 1	Strategy 2
ESG Score	0.0441***	0.0049***
	(2.90)	(2.90)
ESGC Score	0.0424***	0.0047***
	(2.67)	(2.67)
SFDR Score	0.0406***	0.0045***
	(4.02)	(4.02)
SFDR PR	0.0580***	0.0064***
	(3.77)	(3.77)
ESG EW Score	0.0358***	0.0040***
	(2.78)	(2.78)
ESG EW PR	0.0465***	0.0052***
	(2.95)	(2.95)

This table reports annual mean return spreads for portfolio strategies formed on six firm-level sustainability metrics from LSEG over 2002–2023: ESG Score, ESGC Score, SFDR Score, SFDR PR, ESG EW Score, and ESG EW PR. ESG Score and ESGC Score are LSEG proprietary composite ratings. SFDR Score is constructed by aggregating LSEG variables that map into the EU Sustainable Finance Disclosure Regulation Principal Adverse Impact (PAI) framework. "PR" denotes the performance-rank version of a metric, constructed by ranking firms on the underlying indicator inputs prior to aggregation. "ESG EW" denotes an equally weighted variant of the ESG metric based on the same underlying components. For each metric and year, firms are sorted into deciles, where Decile 1 contains the lowest-scoring ("bad") firms and Decile 10 the highest-scoring ("good") firms. Strategy 1 is the long–short spread between Decile 10 and Decile 1 portfolios. Strategy 2 is the spread between a portfolio that excludes Decile 1 (i.e., holds Deciles 2–10) and a portfolio that excludes Decile 10 (i.e., holds Deciles 1–9), each formed using the same sorting metric. Reported values are time-series averages of annual strategy returns over 2002–2023. T-statistics (in parentheses) are based on time-series variation in the annual strategy returns. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

methods and frameworks.

To complement our analysis of returns, we evaluate the risk-adjusted performance of the portfolios using Sharpe ratios. This metric shows how well a strategy turns risk into excess return. It provides a more balanced view of whether portfolios with "good" firms genuinely outperform when volatility is considered. As in the return analysis, we compute the Sharpe ratio for all six sustainability metrics.

Table 4.8 summarizes the Sharpe ratios for the two investment styles emphasizing higher sustainability quality: Strategy 1^g, which invests exclusively in "good" firms, and Strategy 2^g, which excludes only "bad" firms. Sharpe ratios are calculated for one-year returns. Across all metrics, Strategy 1^g achieves higher Sharpe ratios than Strategy 2^g, indicating that concentrated investment in firms with strong sustainability profiles delivers superior risk-adjusted returns.

The SFDR PR score stands out clearly in the Sharpe ratio comparison. Among all sustainability metrics, it produces the highest Sharpe ratio for Strategy 1^g, reaching 0.78, which is the strongest risk-adjusted performance of all the strategies and metrics.

Table 4.7: Annual mean return differences

Score	Strategy 1^g	Strategy 2^g
ESG Score – ESGC Score	0.0024	-0.0001
	(0.76)	(-0.80)
ESG Score – SFDR Score	0.0002	0.0004
	(0.03)	(0.59)
ESG Score – SFDR PR	-0.0033	-0.0012
	(-0.66)	(-1.52)
ESG Score – ESG EW Score	0.0039	0.0005
	(0.81)	(0.52)
ESG Score – ESG EW PR	0.0014	-0.0004
	(0.43)	(-0.30)
ESGC Score – SFDR Score	-0.0022	0.0004
	(-0.51)	(0.69)
ESGC Score – SFDR PR	-0.0056	-0.0011
	(-0.86)	(-1.51)
ESGC Score – ESG EW Score	0.0016	0.0006
	(0.60)	(0.59)
ESGC Score – ESG EW PR	-0.0010	-0.0003
	(-0.25)	(-0.25)
SFDR Score – SFDR PR	-0.0034	-0.0015**
	(-0.58)	(-2.36)
SFDR Score – ESG EW Score	0.0038	0.0001
	(0.88)	(0.20)
SFDR Score – ESG EW PR	0.0012	-0.0008
	(0.23)	(-0.85)
SFDR PR – ESG EW Score	0.0072	0.0017**
	(0.97)	(2.22)
SFDR PR – ESG EW PR	0.0046	0.0008
	(1.00)	(0.86)
ESG EW Score – ESG EW PR	-0.0026	-0.0009
	(-0.59)	(-1.39)

This table reports annual mean return differences for portfolio strategies formed on six firm-level sustainability metrics from LSEG over 2002–2023: ESG Score, ESGC Score, SFDR Score, SFDR PR, ESG EW Score, and ESG EW PR. ESG Score and ESGC Score are LSEG proprietary composite ratings. SFDR Score is constructed by aggregating LSEG variables that map into the EU Sustainable Finance Disclosure Regulation Principal Adverse Impact (PAI) framework. "PR" denotes the performance-rank version of a metric, constructed by ranking firms on the underlying indicator inputs prior to aggregation. "ESG EW" denotes an equally weighted variant of the ESG metric based on the same underlying components. Each row compares two sorting signals and reports the difference in the resulting strategy returns (e.g., the strategy return from sorting on ESG Score minus the strategy return from sorting on ESGC Score). Strategy 1^g uses portfolios invested exclusively in the top decile ("good") firms under each metric. Strategy 2^g uses portfolios that exclude the bottom decile ("bad") firms under each metric. Reported values are time-series averages of annual strategy returns over 2002–2023. T-statistics (in parentheses) are based on time-series variation in the annual strategy returns. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Strategy 2^g also shows its best result under the SFDR PR metric, with a Sharpe ratio of 0.62, outperforming all other metrics. This dual dominance indicates that the SFDR PR classification creates a particularly sharp separation between good and bad firms in terms of their risk–return characteristics. When portfolios fully concentrate on the "good" group, the improvement in risk-adjusted performance is especially pronounced, and even the more moderate negative-screening approach benefits more from the SFDR PR than from any other metric.

Our latest analysis examines the effects of sustainability metrics on long-term firm performance. We calculate the three-year annualized mean returns and Sharpe ratios for each metric for Strategy 1^g and Strategy 2^g.

Table 4.8: Sharpe ratios

Score	Strategy 1^g	Strategy 2^g
ESG Score	0.7099	0.6095
ESGC Score	0.6902	0.6102
SFDR Score	0.7161	0.6029
SFDR PR	0.7798	0.6168
ESG EW Score	0.6809	0.6019
ESG EW PR	0.7228	0.6059

This table reports Sharpe ratios for portfolio strategies formed on six firm-level sustainability metrics from LSEG over 2002–2023: ESG Score, ESGC Score, SFDR Score, SFDR PR, ESG EW Score, and ESG EW PR. ESG Score and ESGC Score are LSEG proprietary composite ratings. SFDR Score is constructed by aggregating LSEG variables that map into the EU Sustainable Finance Disclosure Regulation Principal Adverse Impact (PAI) framework. "PR" denotes the performance-rank version of a metric, constructed by ranking firms on the underlying indicator inputs prior to aggregation. "ESG EW" denotes an equally weighted variant of the ESG metric based on the same underlying components. For each metric and year, firms are sorted into deciles, where Decile 1 contains the lowest-scoring ("bad") firms and Decile 10 the highest-scoring ("good") firms. Strategy 1^g holds the top-decile portfolio for each metric. Strategy 2^g excludes the bottom decile and holds an equal-weighted portfolio of Deciles 2–10. Sharpe ratios are computed from annual strategy returns over 2002–2023.

The three-year annualized results reported in Table 4.9 largely confirm the patterns observed in the shorter-horizon analysis. Strategy 1^g continues to show higher mean returns and higher Sharpe ratios across all metrics, with the SFDR PR metric again producing the strongest outcomes. Importantly, the Sharpe ratio for Strategy 1^g under SFDR PR reaches 1.12, representing the first instance in our results where a Sharpe ratio exceeds 1. This indicates that, over the three-year horizon, the portfolio's excess return more than compensates for its volatility. Strategy 2^g also records its highest Sharpe ratio under the SFDR PR metric at 0.96, consistent with the earlier finding that this metric provides the clearest differentiation in performance between good and bad firms. Overall, the three-year results reinforce the conclusion that portfolios tilted toward firms classified as "good", particularly under the SFDR PR framework, exhibit comparatively stronger risk-adjusted performance.

Table 4.9: Three-year annualized mean returns and sharpe ratios

Score	Mean $S1^g$	SR $S1^g$	Mean $S2^g$	SR $S2^g$
ESG Score	0.0920	0.9621	0.0866	0.9152
ESGC Score	0.0922	0.9568	0.0866	0.9156
SFDR Score	0.0902	0.9339	0.0869	0.9388
SFDR PR	0.0978	1.1192	0.0884	0.9553
ESG EW Score	0.0905	0.9833	0.0865	0.9447
ESG EW PR	0.0917	1.0478	0.0872	0.9572

This table reports three-year annualized mean returns and Sharpe ratios for portfolio strategies formed on six firm-level sustainability metrics from LSEG over 2002–2023: ESG Score, ESGC Score, SFDR Score, SFDR PR, ESG EW Score, and ESG EW PR. ESG Score and ESGC Score are LSEG proprietary composite ratings. SFDR Score is constructed by aggregating LSEG variables that map into the EU Sustainable Finance Disclosure Regulation Principal Adverse Impact (PAI) framework. "PR" denotes the performance-rank version of a metric, constructed by ranking firms on the underlying indicator inputs prior to aggregation. "ESG EW" denotes an equally weighted variant of the ESG metric based on the same underlying components. For each metric and year, firms are sorted into deciles, where Decile 1 contains the lowest-scoring ("bad") firms and Decile 10 the highest-scoring ("good") firms. Strategy 1^g holds the top-decile portfolio. Strategy 2^g excludes the bottom decile and holds an equal-weighted portfolio of Deciles 2–10. Three-year returns are annualized prior to averaging across time; Sharpe ratios are computed from the corresponding annualized strategy returns over 2002–2023.

To assess whether the relative performance differences across sustainability metrics persist over longer horizons, Table 4.10 reports three-year annualized mean return spreads between pairs of sustainability metrics for the two sustainability-oriented strategies. As in the one-year analysis, the majority of pairwise differences are small and statistically insignificant, indicating that portfolio performance is largely robust to the choice of sustainability metric even over extended investment horizons.

Nevertheless, several noteworthy patterns emerge. In particular, comparisons involving the SFDR PR metric continue to stand out. For Strategy 1^g, SFDR PR delivers significantly higher three-year annualized returns relative to both ESG EW Score and ESG EW PR, with spreads of 0.73 p.p. and 0.62 p.p. per year, respectively. These differences are statistically significant at conventional levels and suggest that the adverse-impact-based classification captures dimensions of sustainability that are particularly relevant for long-term return differentiation when portfolios concentrate on high-performing firms.

For the more moderate exclusion-based Strategy 2^g, return spreads remain economically smaller, consistent with earlier findings, but statistically significant differences persist when SFDR PR is compared to ESG-based metrics. Specifically, portfolios excluding low SFDR PR firms outperform those excluding low ESG or ESGC firms by 0.18 p.p. per year. Although modest in magnitude, these effects are statistically meaningful and indicate that the informational advantage of the SFDR PR metric is not confined to short horizons or extreme portfolio tilts.

Table 4.10: Three-year annualized mean return spreads

Score	Strategy 1^g	Strategy 2^g
ESG Score – ESGC Score	-0.0002	0.0000
	(-0.07)	(0.22)
ESG Score – SFDR Score	0.0018	-0.0003
	(0.42)	(-0.39)
ESG Score – SFDR PR	-0.0058	-0.0018***
	(-1.16)	(-2.75)
ESG Score – ESG EW Score	0.0015	0.0001
	(0.47)	(0.09)
ESG Score – ESG EW PR	0.0003	-0.0006
	(0.09)	(-0.42)
ESGC Score – SFDR Score	0.0020	-0.0003
	(0.42)	(-0.40)
ESGC Score – SFDR PR	-0.0057	-0.0018***
	(-0.85)	(-2.73)
ESGC Score – ESG EW Score	0.0016	0.0001
	(0.42)	(0.08)
ESGC Score – ESG EW PR	0.0005	-0.0006
	(0.10)	(-0.43)
SFDR Score – SFDR PR	-0.0077	-0.0015**
	(-1.48)	(-2.57)
SFDR Score – ESG EW Score	-0.0003	0.0004
	(-0.14)	(0.52)
SFDR Score – ESG EW PR	-0.0015	-0.0003
	(-0.33)	(-0.33)
SFDR PR – ESG EW Score	0.0073*	0.0019
	(1.94)	(1.62)
SFDR PR – ESG EW PR	0.0062***	0.0012
	(2.76)	(1.26)
ESG EW Score – ESG EW PR	-0.0011	-0.0007
	(-0.46)	(-1.61)

This table reports three-year annualized mean return spreads for portfolio strategies formed on six firm-level sustainability metrics from LSEG over 2002–2023: ESG Score, ESGC Score, SFDR Score, SFDR PR, ESG EW Score, and ESG EW PR. ESG Score and ESGC Score are LSEG proprietary composite ratings. SFDR Score is constructed by aggregating LSEG variables that map into the EU Sustainable Finance Disclosure Regulation Principal Adverse Impact (PAI) framework. "PR" denotes the performance-rank version of a metric, constructed by ranking firms on the underlying indicator inputs prior to aggregation. "ESG EW" denotes an equally weighted variant of the ESG metric based on the same underlying components. Each row compares two sorting signals and reports the difference in the resulting strategy returns. Strategy 1^g is the long–short spread between the top decile ("good") and bottom decile ("bad") portfolios formed on each metric. Strategy 2^g is the spread between a portfolio that excludes the bottom decile and a portfolio that excludes the top decile for each metric. Three-year returns are annualized prior to averaging across time. Reported t-statistics (in parentheses) are based on time-series variation in the strategy returns. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Taken together, these results provide comprehensive evidence on the relationship between sustainability and financial performance across multiple sustainability metrics, portfolio constructions, and investment horizons. Firms with stronger sustainability profiles consistently achieve higher average returns and superior risk-adjusted performance. This pattern emerges clearly in the decile-based analysis and persists across both concentrated best-in-class portfolios and more moderate exclusion-based strategies.

These findings are consistent with prior evidence documenting positive associations between sustainability and financial performance (e.g., Friede et al. (2015); Khan et al. (2016)), and reinforce the view that sustainability considerations are not inherently at odds with shareholder value maximization. Across all metrics considered, portfolios tilted toward higher-rated firms outperform those concentrated in low-rated firms, indicating that sustainability orientation itself is the primary driver of performance differences.

At the same time, the analysis demonstrates that sustainability metrics are not interchangeable. Although ESG- and SFDR-based classifications are positively correlated and often rank firms similarly, differences in aggregation, normalization, and conceptual emphasis lead to meaningful variation in portfolio composition and performance outcomes. In particular, PR metrics consistently generate a sharper separation between high- and low-performing firms, especially in terms of risk-adjusted returns and long-horizon performance.

Importantly, the results do not suggest that adverse-impact-based metrics uniformly outperform conventional ESG scores. Rather, superior performance arises when sustainability information is combined with performance-rank scaling that emphasizes absolute differences in sustainability performance. This highlights the central role of methodological design in transforming sustainability information into financially relevant investment signals.

These findings also speak to the growing literature on ESG rating disagreement (e.g., Berg et al., 2022). While differences across sustainability metrics do lead to variation in portfolio outcomes, the results indicate that such differences are secondary relative to the broader effect of sustainability orientation itself. Regardless of the specific scoring framework employed, firms classified as sustainability leaders tend to outperform laggards, suggesting that sustainability characteristics retain financial relevance even amid measurement heterogeneity.

4.5 Conclusion

This paper investigates whether and how different sustainability metrics translate into financial performance differences. Using a comprehensive firm-level dataset from LSEG covering the period from 2002 to 2023, we compare conventional ESG scores with adverse-impact-based SFDR metrics and examine the role of aggregation and normalization in shaping sustainability rankings and portfolio outcomes.

Across all specifications, we find robust evidence that sustainability-oriented investment strategies outperform portfolios composed of low-rated firms. Portfolios constructed from top-ranked firms generate significantly higher average returns and superior risk-adjusted performance than those consisting of bottom-ranked firms, irrespective of whether sustainability is measured using traditional ESG scores or SFDR-based indicators. These results suggest that sustainability orientation itself is financially relevant and that incorporating sustainability information into portfolio construction does not come at the expense of performance.

The analysis demonstrates that the choice of sustainability metric materially affects portfolio outcomes. While conventional ESG and ESGC scores, equally weighted ESG metrics, and SFDR-based scores are strongly correlated and largely preserve firms' relative ordering, they produce meaningful differences in return spreads and Sharpe ratios when implemented in investment strategies. Notably, portfolios based on performance-ranked SFDR metrics consistently deliver larger return differentials and higher Sharpe ratios than those based on percentile-ranked ESG metrics, particularly under best-in-class strategies and over longer investment horizons. This suggests that combining adverse-impact indicators with performance-based scaling captures dimensions of sustainability that are less visible in broader ESG scores.

More generally, the results highlight the importance of aggregation and normalization choices in sustainability measurement: performance-rank-based scores emphasize absolute performance gaps rather than relative positioning, revealing differences between average and top-performing firms that percentile-based ESG ratings may obscure. These findings reinforce recent critiques of ESG score construction, indicating that methodological design plays a central role in shaping sustainability assessments and their financial implications.

From an investment perspective, the findings indicate that sustainable investment strategies can achieve superior financial performance, particularly when based on adverse-impact indicators and performance-based rankings. From a regulatory perspective, the results provide empirical support for the relevance of the SFDR framework and its PAI indicators as tools for distinguishing firms along financially meaningful sustainability dimensions.

Several limitations provide avenues for future research. First, while the analysis documents strong associations between sustainability metrics and financial performance, it does not establish causality. Future studies could investigate whether the observed return differentials reflect risk compensation, market mispricing, or operational advantages associated with lower adverse impacts. Second, extending the analysis to other asset classes, international regulatory regimes, or dynamic changes in sustainability performance could further enhance understanding of the mechanisms at work. Finally, future research may explore how investors incorporate adverse-impact information in real time.

Overall, this study shows that sustainability measurement matters for financial performance, that adverse-impact-focused and performance-based metrics offer meaningful advantages over conventional ESG scores, and that sophisticated sustainability integration can enhance both return and risk characteristics of investment portfolios.

Appendix A

Appendix to Chapter 2

A.1 Variable definitions

Table A.1: Definitions and formulas of key variables

Variable	Definition and measurement
Share Variables	
Shares	Number of shares outstanding. For forecast horizons (FY1 to FY5), this is the median analyst consensus forecast for the respective horizon (*TR.NumberofSharesOutstandingMedian(Period=FYi)*). For current or past periods (FY0 to FY-3), this is the number of diluted shares outstanding (*TR.NumberofSharesOutstandingActual(Period=FYi)*, supplemented with *TR.DilutedWghtdAvgShares(Period=FYi)* if missing).
$\Delta\text{Shares}_{FYi}$	Forecasted percentage change in shares for fiscal year i, computed as the change between the forecasted number of shares and the actual shares at the beginning of the period, expressed as a percentage.
Earnings forecast variables	

Table A.1 – continued from previous page

Variable	Definition and measurement
EPS_{FYi}	Earnings per share for fiscal year i. For forecast horizons (FY1 to FY5), this is the median analyst consensus forecast (*TR.EPSMedian(Period=FYi)*). For the current period (FY0), this is the actual reported EPS value (*TR.EPSActValue(Period=FY0)*).
NET_{FYi}	Forecasted net income for horizon Fi, measured as the median analyst consensus. Retrieved from LSEG as *TR.NetProfitMedian_Fi*.
FL_{FYi}	Firm-level EPS for fiscal year i. For forecast horizons (FY1 to FY5), this is computed as analysts' median net profit forecast (*TR.NetProfitMedian(Period=FYi)*)) divided by actual shares outstanding at $FY0$ ($Shares_{FY0}$). For the current period (FY0), it is actual net profit divided by actual shares (*TR.NetProfitActValue(Period=FY0)*) / *TR.NumberofSharesOutstandingActual(Period=FY0)*).
Variables for Implied Cost of Capital Calculation	
B_t	Book value per share. Retrieved from LSEG as *TR.BVPSActValue(Period=FY0)* and, if missing, computed as total equity (*TR.TotalEquity(Period=FY0)*) divided by shares outstanding at $FY0$ ($Shares_{FY0}$).
p_0	Payout ratio. For profit firms (Net Income > 0) calulcated as dividends (*TR.TotalCashDividendsPaid(Period=FY0)*) divided by net income (*TR.NetIncome(Period=FY0)* or if missing *TR.NetIncomeBeforeExtraItems(Period=FY0)*) and for loss firms it is dividends divided by 6% of total assets (*TR.TotalAssetsReported(Period=FY0)*). If dividends are zero, payout ratio is set to 0. The final payout ratio is constrained between 0 and 1.
Firm Characteristics Variables	

Table A.1 – continued from previous page

Variable	Definition and measurement
Growth	Annual sales growth. Calculated as the percentage change in sales using the compustat variable *sale*.
RET	Annual stock return, calculated using the compustat variable *prcc_f* .
Size	Firm size measured as the natural logarithm of market capitalization (*mkvalt*, and if missing *csho* multiplied by *prcc_f*).
LEV	Leverage measured as long-term debt (*dllt*) scaled by total assets (*at*).
RoA	Return on assets measured as EBIT (*ebit*) scaled by total assets (*at*).
FCF	Free cash flow scaled by market equity. Free cash flow is constructed as EBITDA minus capital expenditures and changes in working capital, and is scaled by the firm's market capitalization.
StockComp	Stock-based compensation (*stkco*) scaled by market capitalization.
Forecast Errors of Share Forecasts	
SFE	Scaled forecast error in shares, defined as the difference between actual diluted and forecasted shares outstanding in period i, scaled by diluted shares outstanding in period $FY0$.
SAFE	Scaled absolute forecast error in shares, defined as the absolute difference between actual diluted and forecasted shares outstanding, scaled by diluted shares outstanding in period $FY0$.
Price-scaled Forecast Errors (PFE) and Price-scaled Absolute Forecast Errors (PAFE) of EPS	

Table A.1 – continued from previous page

Variable	Definition and measurement
P(A)FE_EPS	Price-scaled (absolute) forecast error, calculated as actual reported EPS minus forecasted EPS for period i, scaled by the stock price (*prcc_f*).
P(A)FE_Cheat	Price-scaled (absolute) forecast error constructed by converting forecasted net profit for period i into a per-share measure using actual shares outstanding in period i, and subtracting this value from actual reported EPS, scaled by the stock price (*prcc_f*).
P(A)FE_Lagged	Price-scaled (absolute) forecast error constructed by converting forecasted net profit for period i into a per-share measure using lagged actual shares outstanding ($Shares_{FY0}$), and subtracting this value from actual reported EPS, scaled by the stock price (*prcc_f*).
P(A)FE_Forecast	Price-scaled (absolute) forecast error constructed by converting forecasted net profit for period i into a per-share measure using forecasted shares outstanding for period i ($Shares_{FYi}$), and subtracting this value from actual reported EPS, scaled by the stock price (*prcc_f*).

A.2 Implied cost of capital models

Table A.2: Implied cost of capital models

In all models, $E_{t+\tau}$ denotes either the analyst consensus forecast of EPS (EPS_{FYi}) or firm-level earnings (FL_{FYi}) converted to a per-share measure using the actual shares outstanding ($Shares_{FY0}$).

Name	Model/Description
GLS	$P_t = B_t + \Sigma_{\tau=1}^{11} \frac{\mathbb{E}_t[(ROE_{t+\tau} - ICC_{GLS}) \cdot B_{t+\tau-1}]}{(1+ICC_{GLS})^{\tau}} + \frac{\mathbb{E}_t[ROE_{t+12} - ICC_{GLS}) \cdot B_{t+11}]}{(1+ICC_{GLS})^{11} \cdot ICC_{GLS}}$ This model is given by Gebhardt et al. (2001). P_t denotes the stock price as of the estimation date in t, ICC_{GLS} denotes the implied cost of capital (ICC), B_t denotes the book value of equity per share in t and ROE_t is the return on equity in t. B_t is calculated using the clean surplus relation following Hou et al. (2012). ROE_t is calculated by dividing earnings per share (forecasts) E_t by B_{t-1}. For $ROE_{t+\tau}$ up to $\tau = 3$ we use the respective models earnings per share forecast. Afterwards, we assume $ROE_{t+\tau}$ to revert to the historical industry median by $\tau = 11$ (e.g., Hou et al., 2012). The industry median of ROE is derived using 10 years of data while excluding loss firms (e.g., Gebhardt et al., 2001). We expect ROE to be constant after $\tau = 11$.
CT	$P_t = B_t + \sum_{\tau=1}^{5} \frac{\mathbb{E}_t[(ROE_{t+\tau} - ICC_{CT}) \cdot B_{t+\tau-1}]}{(1+ICC_{CT})^{\tau}} + \frac{\mathbb{E}_t[(ROE_{t+5} - ICC_{CT}) \cdot B_{t+4}] \cdot (1+g)}{(1+ICC_{CT})^{5} \cdot (ICC_{CT} - g)}$ This model is given by Claus and Thomas (2001). P_t denotes the stock price as of the estimation date in t, ICC_{CT} denotes the implied cost of capital (ICC), B_t denotes the book value of equity per share in t, ROE_t is the return on equity in t and g is the perpetuity growth rate. B_t is calculated using the clean surplus relation following Hou et al. (2012). ROE_t is calculated by dividing earnings per share (forecasts) E_t by B_{t-1}. g is calculated as the current risk-free rate minus 3% (e.g., Hou et al., 2012).
OJ	$P_t = \frac{\mathbb{E}_t[E_{t+1}] \cdot (g_{st} - (\gamma - 1))}{(R-A)^{-} A^2}, \quad \text{with} \quad A = 0.5((\gamma - 1) \frac{\mathbb{E}_t[E_{t+1}] \cdot payout}{P_t}) \quad g_{st} = \frac{(\mathbb{E}_t[E_{t+2}] - \mathbb{E}_t[E_{t+1}])}{\mathbb{E}_t[E_{t+1}]}$ This model is given by Ohlson and Juettner-Nauroth (2005). P_t denotes the stock price as of the estimation date in t, ICC_{OJ} denotes the implied cost of capital (ICC), $E_{t+\tau}$ is the earnings forecast for $t+\tau$, g_{st} is the short-term growth rate, γ is the perpetual growth rate and *payout* is the current payout ratio. γ is the current risk-free rate minus 3% (e.g., Hou et al., 2012). *payout* is calculated as dividends divided by earnings for profit firms and as dividends divided by 0.06 $\cdot$ total assets for loss firms (e.g., Hou et al. (2012)).
MPEG	$P_t = \frac{\mathbb{E}_t[E_{t+2}] + (ICC_{MPEG} \cdot payout - 1) \cdot \mathbb{E}_t[E_{t+1}]}{ICC_{MPEG}^2}$ This model is given by Easton (2004). P_t denotes the stock price as of the estimation date in t, ICC_{MPEG} denotes the implied cost of capital (ICC) and $E_{t+\tau}$ denotes the earnings per share forecast for $t+\tau$. *payout* is derived as dividends divided by earnings for profit firms and as dividends divided by 0.06 $\cdot$ total assets for loss firms (e.g., Hou et al., 2012).
GG	$P_t = \frac{\mathbb{E}_t[E_{t+1}]}{ICC_{GG}}$ This model is given by Gordon and Gordon (1997). P_t denotes the stock price as of the estimation date in t, ICC_{GG} denotes the implied cost of capital (ICC) and E_{t+1} denotes the earnings per share forecast for $t+1$.

This table gives implied cost of capital (ICC) models that we base our composite ICC on. For simplicity, we drop the firm index i. The composite ICC that we use is derived as the average of these ICC.

Appendix B

Appendix to Chapter 3

B.1 Implied cost of capital models

Table B.1: Implied cost of capital models

In all models, $E_{t+\tau}$ denotes either the analyst consensus forecast of street earnings per share (EPS_{FYi}) or GAAP earnings per share (GPS_{FYi}).

Name	Model/Description
GLS	$P_t = B_t + \Sigma_{\tau=1}^{11} \frac{\mathbb{E}_t[(ROE_{t+\tau} - ICC_{GLS}) \cdot B_{t+\tau-1}]}{(1+ICC_{GLS})^{\tau}} + \frac{\mathbb{E}_t[ROE_{t+12} - ICC_{GLS}) \cdot B_{t+11}]}{(1+ICC_{GLS})^{11} \cdot ICC_{GLS}}$ This model is given by Gebhardt et al. (2001). P_t denotes the stock price as of the estimation date in t, ICC_{GLS} denotes the implied cost of capital (ICC), B_t denotes the book value of equity per share in t and ROE_t is the return on equity in t. B_t is calculated using the clean surplus relation following Hou et al. (2012). ROE_t is calculated by dividing earnings per share (forecasts) E_t by B_{t-1}. For $ROE_{t+\tau}$ up to $\tau = 3$ we use the respective models earnings per share forecast. Afterwards, we assume $ROE_{t+\tau}$ to revert to the historical industry median by $\tau = 11$ (e.g., Hou et al., 2012). The industry median of ROE is derived using 10 years of data while excluding loss firms (e.g., Gebhardt et al., 2001). We expect ROE to be constant after $\tau = 11$.
CT	$P_t = B_t + \sum_{\tau=1}^{5} \frac{\mathbb{E}_t[(ROE_{t+\tau} - ICC_{CT}) \cdot B_{t+\tau-1}]}{(1+ICC_{CT})^{\tau}} + \frac{\mathbb{E}_t[(ROE_{t+5} - ICC_{CT}) \cdot B_{t+4}] \cdot (1+g)}{(1+ICC_{CT})^{5} \cdot (ICC_{CT} - g)}$ This model is given by Claus and Thomas (2001). P_t denotes the stock price as of the estimation date in t, ICC_{CT} denotes the implied cost of capital (ICC), B_t denotes the book value of equity per share in t, ROE_t is the return on equity in t and g is the perpetuity growth rate. B_t is calculated using the clean surplus relation following Hou et al. (2012). ROE_t is calculated by dividing earnings per share (forecasts) E_t by B_{t-1}. g is calculated as the current risk-free rate minus 3% (e.g., Hou et al., 2012).
OJ	$P_t = \frac{\mathbb{E}_t[E_{t+1}] \cdot (g_{st} - (\gamma - 1))}{(R-A)^{-} A^2}$, with $A = 0.5((\gamma - 1) \frac{\mathbb{E}_t[E_{t+1}] \cdot payout}{P_t})$ $g_{st} = \frac{(\mathbb{E}_t[E_{t+2}] - \mathbb{E}_t[E_{t+1}])}{\mathbb{E}_t[E_{t+1}]}$ This model is given by Ohlson and Juettner-Nauroth (2005). P_t denotes the stock price as of the estimation date in t, ICC_{OJ} denotes the implied cost of capital (ICC), $E_{t+\tau}$ is the earnings forecast for $t + \tau$, g_{st} is the short-term growth rate, γ is the perpetual growth rate and *payout* is the current payout ratio. γ is the current risk-free rate minus 3% (e.g., Hou et al., 2012). *payout* is calculated as dividends divided by earnings for profit firms and as dividends divided by $0.06 \cdot$ total assets for loss firms (e.g., Hou et al. (2012)).
MPEG	$P_t = \frac{\mathbb{E}_t[E_{t+2}] + (ICC_{MPEG} \cdot payout - 1) \cdot \mathbb{E}_t[E_{t+1}]}{ICC_{MPEG}^2}$ This model is given by Easton (2004). P_t denotes the stock price as of the estimation date in t, ICC_{MPEG} denotes the implied cost of capital (ICC) and $E_{t+\tau}$ denotes the earnings per share forecast for $t + \tau$. *payout* is derived as dividends divided by earnings for profit firms and as dividends divided by $0.06 \cdot$ total assets for loss firms (e.g., Hou et al., 2012).
GG	$P_t = \frac{\mathbb{E}_t[E_{t+1}]}{ICC_{GG}}$ This model is given by Gordon and Gordon (1997). P_t denotes the stock price as of the estimation date in t, ICC_{GG} denotes the implied cost of capital (ICC) and E_{t+1} denotes the earnings per share forecast for $t + 1$.

This table gives implied cost of capital (ICC) models that we base our composite ICC on. For simplicity, we drop the firm index i. The composite ICC that we use is derived as the average of these ICC.

B.2 Measurement-error-variance diagnostics for expected-return proxies

This appendix summarizes the measurement-error-variance (MEV) framework of Lee et al. (2021), which provides diagnostics for evaluating firm-level expected-return proxies (ERPs), including implied costs of capital (ICCs). The framework yields empirical measures of (i) cross-sectional and (ii) time-series measurement-error variance for a given proxy, scaled in units that facilitate comparisons across ERPs. Lee et al. (2021) propose these diagnostics as a parsimonious way to assess whether an ERP is suitable for ranking firms by expected returns (cross-section) and/or capturing time-variation in expected returns (time series).

Setup and measurement-error decomposition

Let $r_{i,t+1}$ denote the realized return of firm i from t to $t+1$. Let $\mu_{i,t}$ denote firm i's (latent) conditional expected return at time t, and let $\widehat{\mu}_{i,t}$ denote an observed expected-return proxy (ERP), such as an ICC estimate. The proxy can be written as

$$\widehat{\mu}_{i,t} = \mu_{i,t} + \omega_{i,t}, \tag{B.1}$$

where $\omega_{i,t}$ is measurement error. Realized returns satisfy

$$r_{i,t+1} = \mu_{i,t} + \varepsilon_{i,t+1}, \tag{B.2}$$

where $\varepsilon_{i,t+1}$ is the unexpected return (innovation) with $\mathbb{E}[\varepsilon_{i,t+1} \mid \mathcal{F}_t] = 0$.

A key maintained restriction for identification is that the measurement error in the proxy is orthogonal to the unexpected return,

$$\mathrm{Cov}(\omega_{i,t}, \varepsilon_{i,t+1}) = 0, \tag{B.3}$$

so that the proxy's comovement with subsequent realized returns is informative about the proxy's signal content for $\mu_{i,t}$. Lee et al. (2021) show that under this setup, the measurement-error variance can be expressed using moments of $\widehat{\mu}_{i,t}$ and $r_{i,t+1}$.

Scaled measurement-error variance in the cross-section

For a fixed month t, define the cross-sectional variance and covariance operators

$\mathrm{Var}_t(\cdot)$ and $\mathrm{Cov}_t(\cdot,\cdot)$ across firms i. Lee et al. (2021) define the *cross-sectional scaled measurement-error variance* as

$$\mathrm{SVar}_t^{CS}(\omega) \equiv \mathrm{Var}_t(\widehat{\mu}_{i,t}) - 2\,\mathrm{Cov}_t(r_{i,t+1}, \widehat{\mu}_{i,t}). \tag{B.4}$$

Intuition: if $\widehat{\mu}_{i,t}$ is a high-quality proxy for ranking firms by expected returns at time t, it should covary strongly with next period's realized returns in the cross-section, making the second term large and lowering SVar_t^{CS}. Conversely, weak covariance implies larger scaled error variance. (Lower SVar indicates less measurement error in this sense.) Lee et al. (2021) compute SVar_t^{CS} for each month and summarize performance using time-series averages such as

$$\overline{\mathrm{SVar}}^{CS} \equiv \frac{1}{T}\sum_{t=1}^{T} \mathrm{SVar}_t^{CS}(\omega). \tag{B.5}$$

In my implementation, $\widehat{\mu}_{i,t}$ corresponds to an ICC-based proxy (e.g., $ICC_{i,t}^{Street}$, $ICC_{i,t}^{GAAP}$, or their composites), and $r_{i,t+1}$ is the subsequent realized return over the matched horizon.

Scaled measurement-error variance in the time series

To evaluate whether a proxy captures *time-series* variation in expected returns, Lee et al. (2021) define an analogous diagnostic at the firm level. For a fixed firm i, let $\mathrm{Var}_i(\cdot)$ and $\mathrm{Cov}_i(\cdot,\cdot)$ denote time-series variance and covariance over t. The *time-series scaled measurement-error variance* is

$$\mathrm{SVar}_i^{TS}(\omega) \equiv \mathrm{Var}_i(\widehat{\mu}_{i,t}) - 2\,\mathrm{Cov}_i(r_{i,t+1}, \widehat{\mu}_{i,t}). \tag{B.6}$$

The sample-wide summary statistic averages across firms:

$$\overline{\mathrm{SVar}}^{TS} \equiv \frac{1}{N}\sum_{i=1}^{N} \mathrm{SVar}_i^{TS}(\omega). \tag{B.7}$$

A lower $\overline{\mathrm{SVar}}^{TS}$ indicates that, within firms, the proxy comoves more strongly with subsequent realized returns relative to its own variability, consistent with a higher signal-to-noise ratio for time variation in $\mu_{i,t}$.

Implementation details and scaling

Following Lee et al. (2021), I compute SVar_t^{CS} (monthly cross-sections) and SVar_i^{TS}

(firm-level time series) using the realized-return horizon consistent with the proxy construction. In reporting tables, I follow the common convention of multiplying SVar statistics by 100 for readability. Lee et al. (2021) emphasize interpreting SVar comparatively: proxies with smaller scaled measurement-error variances are better suited for the dimension being evaluated (cross-sectional ranking vs. time-series variation).

Appendix C

Appendix to Chapter 4

C.1 Variable definitions

Table C.1: PAI variables used in the SFDR Score construction

Variable	Variable
Accidents_Total	Human_Rights_Controversies
Agrochemical_Products	Human_Rights_Policy
Anti_Personnel_Landmines	Improvement_Tools_Business_Ethics
Biodiversity_Impact_Reduction	Lack_of_a_Supplier_Code_of_Conduct
Board_Gender_Diversity	Land_Environmental_Impact_Reduction
Calculated_Total_Air_Pollutants	Lost_Working_Days
Calculated_Total_Air_Pollutants_2	Non_renewable_energy
Cases_of_insufficient_action_anticorruption_antibribery	OECD_Guidlines
Chemical_or_Biological_Weapons	Ozone_Depleting_Substances
Cluster_Bombs	Policy_Bribery_and_Corruption
CO2_Equivalent_Emissions_Direct_Scope_1	Policy_Child_Labor
CO2_Equivalent_Emissions_Indirect_Scope_2	Policy_Forced_Labor
CO2_Equivalent_Emissions_Indirect_Scope_3	Policy_Supply_Chain_Health_Safety
CO2_Equivalent_Emissions_Scope_1_2_3	Policy_water_efficiency
CO2_Equivalent_Emissions_Total	Renewable_Energy_Purchased
Companies_producing_chemicals	Renewable_Energy_Supply
Contractor_Accidents	Renewable_Energy_Use
Contractor_Lost_working_Days	Resource_Reduction_Policy
Controversies	Salary_Gap
Diversity_and_Opportunity_Controversies	SDG_14_Life_Below_Water
Employee_Accidents	Scope_1_2_3
Employee_Lost_Working_Days	SOx_Emissions
Energy_Purchased_Direct	Supplier_ESG_training
Environmental_Materials_Sourcing	Supply_Chain_Health_Safety_Improvements
Equator_Principles_or_Env_Project_Financing	Supply_Chain_Health_Safety_Training
ESG_Bond_Flag	Total_Energy_Use
Gender_Pay_Gap	TRBC_Fossil_Fuel_Industry_Groups
GHG_Emissions_Intensity_Scope_1_2_3_Paris_Aligned	TRBC_High_Impact_Industry_Groups
GHG_Emissions_Intensity_Scope_1_2_Paris_Aligned	VOC_Emissions
GHG_Emissions_Scope_1_2_3_Paris_Aligned	Waste_Recycling_Ratio
GHG_Emissions_Scope_1_2_Paris_Aligned	Waste_Reduction_Initiatives
Green_Bond_Flag	Water_Pollutant_Emissions
Green_bonds	Water_Recycled_to_Water_Withdrawal_Total
Hazardous_Waste	Water_Withdrawal_Total
Health_Safety_Policy	Whistleblower_Protection
Human_Rights_Contractor	

This table lists all London Stock Exchange Group variables that are relevant for the Principal Adverse Impact statements and are incorporated in the construction of the Sustainable Finance Disclosure Regulation score.

Bibliography

Abarbanell, J. S. and B. J. Bushee (1997). Fundamental analysis, future earnings, and stock prices. *Journal of Accounting Research 35*(1), 1.

Abouarab, R., T. Mishra, and S. Wolfe (2025). Does the eu sustainable finance disclosure regulation mitigate greenwashing? *The European Journal of Finance 31*(8), 957–989.

Albuquerque, R., Y. Koskinen, S. Yang, and C. Zhang (2020). Resiliency of environmental and social stocks: An analysis of the exogenous covid-19 market crash. *The Review of Corporate Finance Studies 9*(3), 593–621.

Almeida, H., V. Fos, and M. Kronlund (2016). The real effects of share repurchases. *Journal of Financial Economics 119*(1), 168–185.

Amiraslani, H., K. V. Lins, H. Servaes, and A. Tamayo (2023). Trust, social capital, and the bond market benefits of esg performance. *Review of Accounting Studies 28*(2), 421–462.

Asquith, P. and D. W. Mullins Jr (1986). Equity issues and offering dilution. *Journal of Financial Economics 15*(1-2), 61–89.

Bae, J., X. Yang, and M.-I. Kim (2021). Esg and stock price crash risk: role of financial constraints. *Asia-Pacific Journal of Financial Studies 50*(5), 556–581.

Baker, M. and J. Wurgler (2002). Market timing and capital structure. *The Journal of Finance 57*(1), 1–32.

Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics 9*(1), 3–18.

Becker, M. G., F. Martin, and A. Walter (2022). The power of esg transparency: The effect of the new sfdr sustainability labels on mutual funds and individual investors. *Finance Research Letters 47*(Part B), 102708.

Bens, D. A., V. Nagar, D. J. Skinner, and M. F. Wong (2003). Employee stock options, eps dilution, and stock repurchases. *Journal of Accounting and Economics 36*(1-3), 51–90.

Benuzzi, M., K. Bax, S. Paterlini, and E. Taufer (2025). Chasing esg performance: How methodologies shape outcomes. *International Review of Financial Analysis*, 104239.

Berg, F., J. F. Kölbel, and R. Rigobon (2022). Aggregate confusion: The divergence of esg ratings. *Review of Finance 26*(6), 1315–1344.

Bhattacharya, N., E. L. Black, T. E. Christensen, and C. R. Larson (2003). Assessing the relative informativeness and permanence of pro forma earnings and gaap operating earnings. *Journal of Accounting and Economics 36*(1-3), 285–319.

Bradshaw, M. T., T. E. Christensen, K. H. Gee, and B. C. Whipple (2018). Analysts' gaap earnings forecasts and their implications for accounting research. *Journal of Accounting and Economics 66*(1), 46–66.

Bradshaw, M. T., R. Holstein III, J. Tucker, and J. Zhou (2025). Analysts' explicit forecasts of the number of shares outstanding. *Available at SSRN 5864842*.

Bradshaw, M. T. and R. G. Sloan (2002). Gaap versus the street: An empirical assessment of two alternative definitions of earnings. *Journal of Accounting Research 40*(1), 41–66.

Broadstock, D. C., K. Chan, L. T. Cheng, and X. Wang (2021). The role of esg performance during times of financial crisis: Evidence from covid-19 in china. *Finance Research Letters 38*, 101716.

Brown, L. D. (2001). A temporal analysis of earnings surprises: Profits versus losses. *Journal of Accounting Research 39*(2), 221–241.

Brown, L. D. and K. Sivakumar (2003). Comparing the value relevance of two operating income measures. *Review of Accounting Studies 8*(4), 561–572.

Busch, D. (2023). Eu sustainable finance disclosure regulation. *Capital Markets Law Journal 18*(3), 303–328.

Callen, J. L. and M. R. Lyle (2020). The term structure of implied costs of equity capital. *Review of Accounting Studies 25*(1), 342–404.

CFA Institute, Global Sustainable Investment Alliance, and Principles for Responsible Investment (2023). Definitions for responsible investment approaches. https://www.gsi-alliance.org/wp-content/uploads/2023/10/ESG-Terminology-Report_Online.pdf.

Chatterji, A. K., R. Durand, D. I. Levine, and S. Touboul (2016). Do ratings of firms converge? implications for managers, investors and strategy researchers. *Strategic Management Journal 37*(8), 1597–1614.

Christensen, T., M. Lewis-Western, C. Rabe, and H. J. Wang (2025). Do we really know what we think we know about investors' reliance on gaap and non-gaap earnings? Technical report, Working Paper.

Claus, J. and J. Thomas (2001). Equity premia as low as three percent? evidence from analysts' earnings forecasts for domestic and international stock markets. *The Journal of Finance 56*(5), 1629–1666.

Cohen, D. A., R. N. Hann, and M. Ogneva (2007). Another look at gaap versus the street: an empirical assessment of measurement error bias. *Review of Accounting Studies 12*(2), 271–303.

Cornett, M. M., O. Erhemjamts, and H. Tehranian (2016). Greed or good deeds: An examination of the relation between corporate social responsibility and the financial performance of us commercial banks around the financial crisis. *Journal of Banking & Finance 70*, 137–159.

Dechow, P. M., S. A. Richardson, and I. Tuna (2003). Why are earnings kinky? an examination of the earnings management explanation. *Review of Accounting Studies 8*, 355–384.

Demers, E., J. Hendrikse, P. Joos, and B. Lev (2021). Esg did not immunize stocks during the covid-19 crisis, but investments in intangible assets did. *Journal of Business Finance & Accounting 48*(3-4), 433–462.

Diether, K. B., C. J. Malloy, and A. Scherbina (2002). Differences of opinion and the cross section of stock returns. *The Journal of Finance 57*(5), 2113–2141.

Doyle, J., R. Lundholm, and M. Soliman (2003). The predictive value of expenses excluded from pro forma earnings. *Review of Accounting Studies 8*(2-3), 145–174.

Easton, P. D. (2004). Pe ratios, peg ratios, and estimating the implied expected rate of return on equity capital. *The Accounting Review 79*(1), 73–95.

Easton, P. D. and S. J. Monahan (2005). An evaluation of accounting-based measures of expected returns. *The Accounting Review 80*(2), 501–538.

Easton, P. D. and G. A. Sommers (2007). Effect of analysts' optimism on estimates of the expected rate of return implied by earnings forecasts. *Journal of Accounting Research 45*(5), 983–1015.

Eccles, R. G., I. Ioannou, and G. Serafeim (2014). The impact of corporate sustainability on organizational processes and performance. *Management Science 60*(11), 2835–2857.

Eisfeldt, A. L. and D. Papanikolaou (2013). Organization capital and the cross-section of expected returns. *The Journal of Finance 68*(4), 1365–1406.

El Ghoul, S., O. Guedhami, H. Kim, and K. Park (2018). Corporate environmental responsibility and the cost of capital: International evidence. *Journal of Business Ethics 149*(2), 335–361.

El Ghoul, S., O. Guedhami, C. C. Kwok, and D. R. Mishra (2011). Does corporate social responsibility affect the cost of capital? *Journal of Banking & Finance 35*(9), 2388–2406.

El Ghoul, S., O. Guedhami, C. C. Kwok, and D. R. Mishra (2011). Does corporate social responsibility affect the cost of capital? *Journal of Banking & Finance 35*(9), 2388–2406.

Engelhardt, N., J. Ekkenga, and P. Posch (2021). Esg ratings and stock performance during the covid-19 crisis. *Sustainability 13*(13), 7133.

European Commission (2022). Commission Delegated Regulation (EU) 2022/1288 of 6 April 2022 supplementing Regulation (EU) 2019/2088 with regard to regulatory technical standards (SFDR RTS). `https://eur-lex.europa.eu/eli/reg_del/2022/1288/oj/eng`.

European Parliament and the Council of the European Union (2019). Regulation (EU) 2019/2088 of 27 November 2019 on sustainability-related disclosures in the financial services sector (SFDR). `https://eur-lex.europa.eu/eli/reg/2019/2088/oj/eng`.

Fama, E. F. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy 81*(3), 607–636.

Fama, E. F. and K. R. French (1995). Size and book-to-market factors in earnings and returns. *Journal of Finance 50*(1), 131–155.

Fama, E. F. and K. R. French (2005). Financing decisions: who issues stock? *Journal of Financial Economics 76*(3), 549–582.

Francis, J. and L. Soffer (1997). The relative informativeness of analysts' stock recommendations and earnings forecast revisions. *Journal of Accounting Research 35*(2), 193–211.

Friede, G., T. Busch, and A. Bassen (2015). Esg and financial performance: aggregated evidence from more than 2000 empirical studies. *Journal of Sustainable Finance & Investment 5*(4), 210–233.

Gebhardt, W. R., C. M. C. Lee, and B. Swaminathan (2001). Toward an implied cost of capital. *Journal of Accounting Research 39*(1), 135–176.

Gibson Brandon, R., P. Krueger, and P. S. Schmidt (2021). Esg rating disagreement and stock returns. *Financial Analysts Journal 77*(4), 104–127.

Giese, G., L.-E. Lee, D. Melas, Z. Nagy, L. Nishikawa, et al. (2019). Foundations of esg investing: How esg affects equity valuation, risk, and performance. *The Journal of Portfolio Management 45*(5), 69–83.

Gillan, S. L., A. Koch, and L. T. Starks (2021). Firms and social responsibility: A review of esg and csr research in corporate finance. *Journal of Corporate Finance 66*, 101889.

Gode, D. and P. Mohanram (2003). Inferring the cost of capital using the ohlson–juettner model. *Review of Accounting Studies 8*, 399–431.

Godfrey, P. C., C. B. Merrill, and J. M. Hansen (2009). The relationship between corporate social responsibility and shareholder value: An empirical test of the risk management hypothesis. *Strategic Management Journal 30*(4), 425–445.

Gordon, J. R. and M. J. Gordon (1997). The finite horizon expected return model. *Financial Analysts Journal 53*(3), 52–61.

Graham, J. R., C. R. Harvey, and S. Rajgopal (2005). The economic implications of corporate financial reporting. *Journal of Accounting and Economics 40*(1-3), 3–73.

Guay, W., S. Kothari, and S. Shu (2011). Properties of implied cost of capital using analysts' forecasts. *Australian Journal of Management 36*(2), 125–149.

Guerrero, S. and J. P. Viteri (2025). What are environmental, social, and governance scores measuring? the role of outcome and impact indicators in esg scores. *Finance Research Letters 72*, 106529.

Gupta, K. (2018). Environmental sustainability and implied cost of equity: International evidence. *Journal of Business Ethics 147*(2), 343–365.

Hoepner, A. G., I. Oikonomou, Z. Sautner, L. T. Starks, and X. Y. Zhou (2024). Esg shareholder engagement and downside risk. *Review of Finance 28*(2), 483–510.

Hong, H. and J. D. Kubik (2003). Analyzing the analysts: Career concerns and biased earnings forecasts. *The Journal of Finance 58*(1), 313–351.

Hou, K., M. A. Van Dijk, and Y. Zhang (2012). The implied cost of capital: A new approach. *Journal of Accounting and Economics 53*(3), 504–526.

Hribar, P., N. T. Jenkins, and W. B. Johnson (2006). Stock repurchases as an earnings management device. *Journal of Accounting and Economics 41*(1-2), 3–27.

Ikenberry, D., J. Lakonishok, and T. Vermaelen (1995). Market underreaction to open market share repurchases. *Journal of Financial Economics 39*(2-3), 181–208.

Jo, H. and H. Na (2012). Does csr reduce firm risk? evidence from controversial industry sectors. *Journal of Business Ethics 110*(4), 441–456.

Johnson, W. B. and W. C. Schwartz Jr (2005). Are investors misled by "pro forma" earnings? *Contemporary Accounting Research 22*(4), 915–963.

Jolls, C. (1998). Stock repurchases and incentive compensation. *NBER Working Paper 6467*.

Kang, N. and R. Sadka (2017). Implied cost of capital in the cross-section of stocks. *Available at SSRN 2600147*.

Kaplan, Z., N. T. Marshall, J. D. Mathis, and H. I. Wang (2024). Do two wrongs make a right? strategically forecasting eps through inaccurate share forecasts. *Working Paper*.

Khan, M., G. Serafeim, and A. Yoon (2016). Corporate sustainability: First evidence on materiality. *The Accounting Review 91*(6), 1697–1724.

Kolev, K., C. A. Marquardt, and S. E. McVay (2008). Sec scrutiny and the evolution of non-gaap reporting. *The Accounting Review 83*(1), 157–184.

Landsman, W. R., B. L. Miller, and S. Yeh (2007). Implications of components of income excluded from pro forma earnings for future profitability and equity valuation. *Journal of Business Finance & Accounting 34*(3-4), 650–675.

Lee, C. M., E. C. So, and C. C. Wang (2021). Evaluating firm-level expected-return proxies: implications for estimating treatment effects. *The Review of Financial Studies 34*(4), 1907–1951.

Lev, B. and P. Zarowin (1999). The boundaries of financial reporting and how to extend them. *Journal of Accounting Research 37*(2), 353–385.

Li, K. K. and P. Mohanram (2014). Evaluating cross-sectional forecasting models for implied cost of capital. *Review of Accounting Studies 19*(3), 1152–1185.

Li, Y., D. T. Ng, and B. Swaminathan (2013). Predicting market returns using aggregate implied cost of capital. *Journal of Financial Economics 110*(2), 419–436.

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

Livnat, J. and R. R. Mendenhall (2006). Comparing the post–earnings announcement drift for surprises calculated from analyst and time series forecasts. *Journal of Accounting Research 44*(1), 177–205.

Loughran, T. and J. R. Ritter (1995). The new issues puzzle. *The Journal of Finance 50*(1), 23–51.

Mikhail, M. B., B. R. Walther, and R. H. Willis (1997). Do security analysts improve their performance with experience? *Journal of Accounting Research 35*, 131–157.

Miller, M. H. and F. Modigliani (1961). Dividend policy, growth, and the valuation of shares. *the Journal of Business 34*(4), 411–433.

Myers, S. C. and N. S. Majluf (1984). Corporate financing and investment decisions when firms have information that investors do not have. *Journal of Financial Economics 13*(2), 187–221.

Ohlson, J. A. and B. E. Juettner-Nauroth (2005). Expected eps and eps growth as determinants of value. *Review of Accounting Studies 10*(2), 349–365.

Oikonomou, I., C. Brooks, and S. Pavelin (2012). The impact of corporate social performance on financial risk and utility: A longitudinal analysis. *Financial Management 41*(2), 483–515.

Pedersen, L. H., S. Fitzgibbons, and L. Pomorski (2021). Responsible investing: The esg-efficient frontier. *Journal of Financial Economics 142*(2), 572–597.

Petersen, M. A. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies 22*(1), 435–480.

Prol, J. L. and K. Kim (2022). Risk-return performance of optimized esg equity portfolios in the nyse. *Finance Research Letters 50*, 103312.

Scheitza, L. and T. Busch (2024). Sfdr article 9: Is it all about impact? *Finance Research Letters 62*, 105179.

Vermaelen, T. (1981). Common stock repurchases and market signalling: An empirical study. *Journal of Financial Economics 9*(2), 139–183.

Welch, I. (2000). Views of financial economists on the equity premium and on professional controversies. *The Journal of Business 73*(4), 501–537.

Whelan, T., U. Atz, T. Van Holt, and C. Clark (2021). Esg and financial performance. *Uncovering the Relationship by Aggregating Evidence from 1*(2015-2020), 10.

Eidesstattliche Erklärung

Hiermit versichere ich an Eides Statt, dass ich die vorgelegte Arbeit selbstständig und ohne die Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus anderen Quellen direkt oder indirekt übernommenen Aussagen, Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet. Bei der Auswahl und Auswertung folgenden Materials haben mir die nachstehend aufgeführten Personen in der jeweils beschriebenen Weise entgeltlich/unentgeltlich geholfen:

Prof. Dr. Dieter Hess und Luca Conrads (gemeinsame Projektarbeit für die Inhalte in Kapitel 2 und Kapitel 4).

Weitere Personen, neben den ggf. in der Einleitung der Arbeit aufgeführten Koautorinnen und Koautoren, waren an der inhaltlich-materiellen Erstellung der vorliegenden Arbeit nicht beteiligt. Insbesondere habe ich hierfür nicht die entgeltliche Hilfe von Vermittlungs- bzw. Beratungsdiensten in Anspruch genommen. Niemand hat von mir unmittelbar oder mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen.

Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

Ich versichere, dass ich nach bestem Wissen die reine Wahrheit gesagt und nichts verschwiegen habe.

Ich versichere, dass die eingereichte elektronische Fassung der eingereichten Druckfassung vollständig entspricht.

Die Strafbarkeit einer falschen eidesstattlichen Versicherung ist mir bekannt, namentlich die Strafandrohung gemäß § 156 StGB bis zu drei Jahren Freiheitsstrafe oder Geldstrafe bei vorsätzlicher Begehung der Tat bzw. gemäß § 161 Abs. 1 StGB bis zu einem Jahr Freiheitsstrafe oder Geldstrafe bei fahrlässiger Begehung.

Ort, Datum | Unterschrift