Probability & Partners,
Amsterdam Tech
Manos Ieronymakis
June 2024

Navigating the ESG Landscape:

A Comparative Analysis of Slow and Fast Scores and Their Applications

---

Data Science And Leadership Thesis

Table of Contents

Project summary

This project explores the integration of high-frequency (fast) and traditional (slow) Environmental, Social, and Governance (ESG) scores to assess S&P 500 companies. The analysis focuses on the relationship between these two types of scores from various sources, measuring how well fast ESG scores can predict traditional ones. Additionally, the research introduces a novel approach to create a greenwashing indicator derived from the ESG data. By combining both high frequency and traditional ESG scores, this study offers a current and thorough perspective on ESG performance, ultimately aiming to enhance how investors evaluate companies in the S&P 500.

1. Introduction

In today's fast-moving financial world, something big is happening: Environmental, Social, and Governance (ESG) criteria are taking over. They are not just a side note anymore; they are front and centre, catching the eye of big investors and everyday traders. This shift is changing the game for how people look at companies, putting their long-term sustainability and impact on society under the microscope. This thesis primarily aims to explore this exciting new landscape, exploring the connection between two types of ESG scores – the slow and the fast. In this paper, my goal is to guide readers through the complex relationship between these scores and how these scores affect companies' performance and overall financial sentiment.

Before I describe the specifics of slow and fast ESG scores, it is important to understand what ESG component includes. Environmental factors involve a company's impact on the planet, including its carbon footprint, resource usage, and adherence to environmental regulations. Social considerations revolve around the company's relationships with its employees, communities, and broader society, covering issues such as labour practices, diversity, and community engagement. Governance criteria relate to the structure and oversight of the company, focusing on aspects like board diversity, executive compensation and transparency in decision-making processes.

Slow ESG scores, which in this research are provided by Refinitiv Ltd, represent a traditional method of evaluating companies' ESG performance. These scores typically rely on annual or periodic disclosures provided by companies themselves, as well as data from regulatory filings and third-party assessments. While slow ESG scores offer a snapshot of a company's ESG performance over time, they may face limitations such as data lag and reliance on self-reported information.

In contrast, fast ESG scores, illustrated by MarketPsych PTE Ltd, leverage advanced technologies such as Natural Language Processing (NLP) and machine learning to provide real-time or near-real-time insights into companies' ESG performance. These scores are derived from the analysis of news articles, social media posts, and other textual sources, enabling a more dynamic and responsive assessment of ESG-related events and trends.

This research has significant implications for Probability & Partners B.V., a renowned risk management consultancy in the financial sector. The project is supported by a diverse network including ESG experts, financial institutions such as Probability & Partners, and key data providers like Refinitiv and MarketPsych. Over the next six months, I will benefit from the guidance of two senior risk management consultants who specialize in ESG and risk management. Additionally, I am supervised by a partner at Probability & Partners who is also a professor at Vrije Universiteit Amsterdam, with deep expertise in ESG issues.

This study's examination of the dynamics between slow and fast ESG scores provides Probability & Partners with actionable insights and influences my broader stakeholder network. Both retail and institutional traders gain from these findings by acquiring new perspectives on S&P 500 companies, potentially improving their trading strategies and ESG comprehension. Furthermore, Probability & Partners secures a market advantage by integrating innovative services based on leading-edge academic research. Academic institutions and research organizations also benefit from the use of novel data, which advances academic dialogue and encourages further research.

On a personal level, this thesis represents a journey of growth and exploration for me. With an interest in stock market and ESG topics, this project not only aligns with my career aspirations, but also presents an opportunity to expand my expertise in this domain. It is about pushing my boundaries of knowledge and accepting the challenges of diverse research areas.

Additionally, this research taps into the broader trend of sentiment data analysis in the financial landscape. As sentiment data becomes increasingly influential in decision-making processes, there is a remarkable opportunity to develop models that translate this data into actionable insights.

In summary, this thesis constitutes a deep dive into the intersection of finance, sustainability, and technology. It aims to comprehend the past, present, and future of ESG investing, uncovering the underlying patterns and correlations that drive financial markets. It leverages data science and leadership skills to make a meaningful impact on the world. Above all, it embraces the journey of discovery and growth, both personally and professionally, navigating the complexities of the financial universe.

2. Terms of reference/objectives and literature review/desktop research

This chapter presents the goals and provides a review of the existing literature related to this project. The research statement focuses on understanding the relationship between self-reported (slow) ESG scores and high-frequency (fast) ESG scores among S&P 500 companies. The research questions aim to explore the predictive power of fast ESG data over slow ESG data and identify discrepancies that might indicate greenwashing. The chapter further details various objectives tailored for researchers, the project, and stakeholders interested in ESG scores.

2.1 Terms of reference/objectives

2.1.1 Research statement

The research statement is driven by the need to investigate the patterns between different types of ESG scores and detect potential greenwashing practices. By focusing on both self-reported and market sentiment-based ESG scores, this study aims to provide a deeper understanding of ESG metrics and their implications for S&P 500 companies.

2.1.2 Research questions

The main research questions guiding this study are: How do slow ESG scores correlate with fast ESG scores for companies in the S&P 500 Index? Can fast ESG scores predict changes in slow ESG scores over time? What patterns and trends emerge from the comparative analysis of slow and fast ESG scores? How can a machine learning model be developed to measure the predictive power of fast ESG data over slow ESG data? Finally, what methodology can be used to detect potential greenwashing practices at the company level?

2.1.3 Objectives

The objectives of this project are multifaceted, addressing the needs of the researcher which is me and stakeholders interested in ESG scores which are Refinitiv and MarketPsych.

For the researcher, the goals include developing a predictive model using ARIMA to examine the relationship between slow and fast ESG scores, conducting exploratory data analysis to uncover patterns and correlations, and implementing a methodology for detecting potential greenwashing practises.

For the project, the primary objectives go beyond numerical predictions, aiming to identify trends and potential discrepancies between slow and fast ESG scores. Additionally, the project seeks to assess the impact of ESG scores on the financial performance and sustainability practices of S&P 500 companies.

For stakeholders, the project aims to provide actionable insights that aid in making informed decisions regarding ESG investments and corporate sustainability practices, and to offer a methodology for identifying potential greenwashing to promote transparency and accountability. Specifically, both MarketPsych and Refinitiv are keen to explore the correlation between their data to better understand what they offer to their clients and assess the reliability of their data sources.

2.2 Literature review

In this literature review, I focus on ESG data, predictive modelling, and greenwashing detection to ground my research in existing knowledge. I used handbooks from Refinitiv and MarketPsych, which provide detailed insights into how each provider constructs and delivers ESG scores. To build my predictive models, I reviewed methodologies from similar studies that use machine learning to forecast ESG ratings, aligning with my goal of using data science to analyse ESG scores from different sources and frequencies. For greenwashing detection, I explored established strategies that utilize sentiment data to identify companies prone to greenwashing. The company that inspired me the most in developing a greenwashing indicator is called Covalence, which specializes in greenwashing applications.

3. Methodology

This project examines two types of ESG scores among S&P 500 companies. The primary aim is to determine whether these ESG scores are correlated and how they can illustrate the general financial environment of these corporations.

Project steps:

1. **Exploratory Data Analysis**: Initially, I will review the two data sources to identify similarities and differences in the fast and slow ESG data. The objective is to understand how S&P 500 companies fare in their ESG commitments based on self-reporting versus societal and public perception. This analysis will pave the way for identifying correlations and patterns in how companies are rated on ESG metrics. The goal is to form general assumptions and categorize companies by size (market capitalization) and industry sector.
2. **Model Development**: I will employ autoregressive models to measure the predictive power of ESG scores. This objective is crucial because slow scores are susceptible to changes post-publication, unlike fast scores, which do not change after publication. Using ARIMA models, I will attempt to predict fluctuations in slow scores over time.
3. **Greenwashing Indicator Development**: Additionally, the project will expand to include a greenwashing indicator. Using the same data employed in the modelling, I will attempt to create an implementation for measuring greenwashing at the company level for some of the S&P 500 firms.

Research methodology:

1. **Data collection/Description**: I will collect data from two sources, Refinitiv and MarketPsych. It is essential to provide an analytical description of how these ESG scores are generated, to aid the reader's understanding of their origins and complexity.
2. **Data cleaning and analysis**: This phase involves cleaning the data and conducting an in-depth analysis to assess the data's reliability and explore potential correlations within the S&P 500 Index. At the same time, I will research to gather market capitalization and sector data for all companies to group them for potential pattern analysis.
3. **Model development and greenwashing indicator development**: I will model the results using ARIMA models to predict the slow ESG scores based on fast ones. Simultaneously, using the same data, I will build a greenwashing indicator that compares the two data sources' ESG scores to discern if companies are prone to greenwashing.
4. **Comparative analysis**: I will apply the greenwashing indicator to four companies from the S&P 500 that meet some predefined correlation criteria to derive insights into which years they may have greenwashed their results.

Cycles for action research:

1. **Idea formation**: Initially, I will refine the project idea through one-on-one coaching with an ESG expert and partner in my working environment. This ensures that the project's objectives are valuable for the researcher, my organization, and our clients who provide the data.

2. **Collaborative review**: I will continue these discussions in collaborative sessions with my employer and the data providers (MarketPsych and Refinitiv) to ensure that the project's objectives align with their interests and receive additional guidance on focusing the project.
3. **University Mentorship**: After gathering different perspectives and achieving agreement on the project's objectives, I will discuss these ideas with my university mentor to ensure they align with the project's goals and to gain invaluable insights for further development.
4. **Exploratory data analysis and visualisation**: I will then begin exploring the dataset and conducting exploratory data analysis and visualization. This analytical approach will help understand the dataset details and identify the best methodologies to predict my project's objectives.
5. **Iterative testing**: I plan to repeat this process with my organizational mentor, which will be an ongoing process of testing, evaluation, and refinement to ensure the project is as effective as possible and achieves the best outcomes.

Nevertheless, there are certain limitations to my methodology. ESG scores provided by different data sources can vary significantly, which may affect the applicability of the findings to other sets of ESG scores. Additionally, the greenwashing indicator is limited to data only up to the year 2022. Furthermore, a company that I consulted for greenwashing methodology refused to share their results, preventing a comparative analysis to determine if different methodologies yield consistent findings on ESG score misalignments. Finally the action research practise generally works well; however, it is time-consuming in my case because I need to communicate with many individuals from different organizations and repeatedly present my methodologies and findings. This process slows me down and sometimes leads to confusion, as the results are subjective and can be interpreted differently by various teams.

3.1 Slow ESG data construction

In this section, I will detail the methodology used to calculate slow ESG scores as provided by Refinitiv. Unlike fast scores, which are designed to quickly respond to market sentiment and news, Refinitiv's slow scores rely on comprehensive data collection and rigorous analytics, offering a stable and in-depth reflection of a company's ESG performance over time.

Refinitiv's ESG scoring framework is built from data sourced from various public disclosures, including annual reports, CSR[1] reports, and information released by NGOs[2]. This collection covers over 630 ESG metrics (Refinitiv, 2022). From this vast dataset, approximately 186 key metrics are distilled, chosen for their comparability and material relevance across industries, providing the foundation for the assessment and scoring process. These metrics span ten categories, such as emissions, human rights, and product responsibility, ensuring a comprehensive evaluation of a company's ESG practices (Refinitiv, 2022).

[1] CSR reports: Comprehensive documents published by companies outlining their Corporate Social Responsibility initiatives and performance.
[2] Information released by NGOs: Data and reports provided by Non-Governmental Organizations regarding various social, environmental, and governance issues.

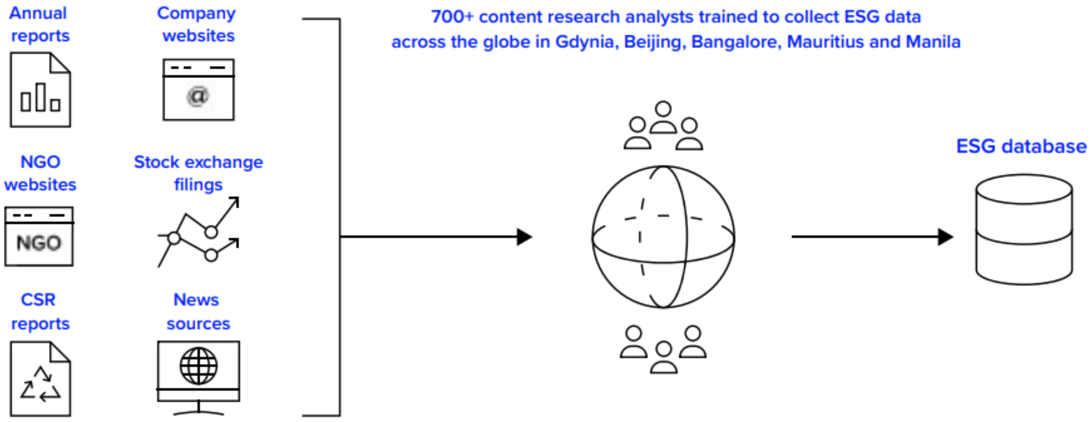

Figure 1: ESG data collection process

Generating a slow ESG score starts with treating the raw data points, which are classified as either Boolean or numeric. Boolean data are converted into numeric scores, reflecting the presence or absence of specific policies and practices, while numeric data are ranked percentile-wise to enable comparisons across companies in similar industries. This normalization process adjusts for varying scales and operational scopes (Refinitiv, 2022).

Next, the Materiality Matrix assigns different weights to the ESG categories based on their importance in specific industries. The model recognizes that ESG issues do not impact all sectors equally. For instance, emissions are more significant in manufacturing than in services. These weights are derived from the industry's median values and data transparency, emphasizing disclosure quality in the scoring process (LSEG, 2022).

Pillars	Catagories	Themes	Data points	Weight method
Environmental	Emmission	Emissions	TR.AnalyticCO2	Quant industry median
		Waste	TR.AnalyticTotalWaste	Quant industry median
		Biodiversity*		
		Environmental management systems*		
	Innovation	Product innovation	TR.EnvProducts	Transparency weights
		Green revenues, research and development (R&D) and capital expenditures (CapEx)	TR.AnalyticEnvRD	Quant industry median
	Resource use	Water	TR.AnalyticWaterUse	Quant industry median
		Energy	TR.AnalyticEnergyUse	Quant industry median
		Sustainable packaging*		
		Environmental supply chain*		
Social	Community	Equally important to all industry groups, hence a median weight of five is assigned to all		Equally important to all industry groups
	Human rights	Human rights	TR.PolicyHumanRights	Transparency weights
	Product responsibility	Responsible marketing	TR.PolicyResponsibleMarketing	Transparency weights
		Product quality	TR.ProductQualityMonitoring	Transparency weights
		Data privacy	TR.PolicyDataPrivacy	Transparency weights
	Workforce	Diversity and inclusion	TR.WomenEmployees	Quant industry median
		Career development and training	TR.AvgTrainingHours	Transparency weights
		Working conditions	TR.TradeUnionRep	Quant industry median
		Health and safety	TR.AnalyticLostDays	Transparency weights
Governance	CSR strategy	CSR strategy	Data points in governance category and governance pillar	Count of data points in each governance category/all data points in governance pillar
		ESG reporting and transparency		
	Management	Structure (independence, diversity, committees)	Data points in governance category and governance pillar	Count of data points in each governance category/all data points in governance pillar
		Compensation		
	Shareholders	Shareholder rights	Data points in governance category and governance pillar	Count of data points in each governance category/all data points in governance pillar
		Takeover defenses		

Figure 2: ESG materiality matrix categories

Once the category weights are set, each category's scores are calculated by aggregating and adjusting the data points according to their respective weights. These scores are then grouped into three main ESG pillars: Environmental, Social, and Governance. Each pillar's score is a sum of its category scores, which are normalized to produce a final score on a 0-100 scale, where a higher score indicates better ESG performance.

The scoring process incorporates a controversies score to account for real-time events that may affect a company's reputation and stakeholder trust. The controversies score adjusts the overall ESG score based on a company's involvement in adverse ESG incidents. This adjustment ensures the scores reflect the most current state of a company's ESG performance, including any recent negative events (Refinitiv, 2022).

Finally, the ESG Combined (ESGC) score blends the overall ESG score with the controversies score, providing a comprehensive view of a company's ESG profile that considers proactive initiatives and potential risks tied to negative events. This score gives investors and stakeholders a nuanced perspective on a company's ESG status (Refinitiv, 2022).

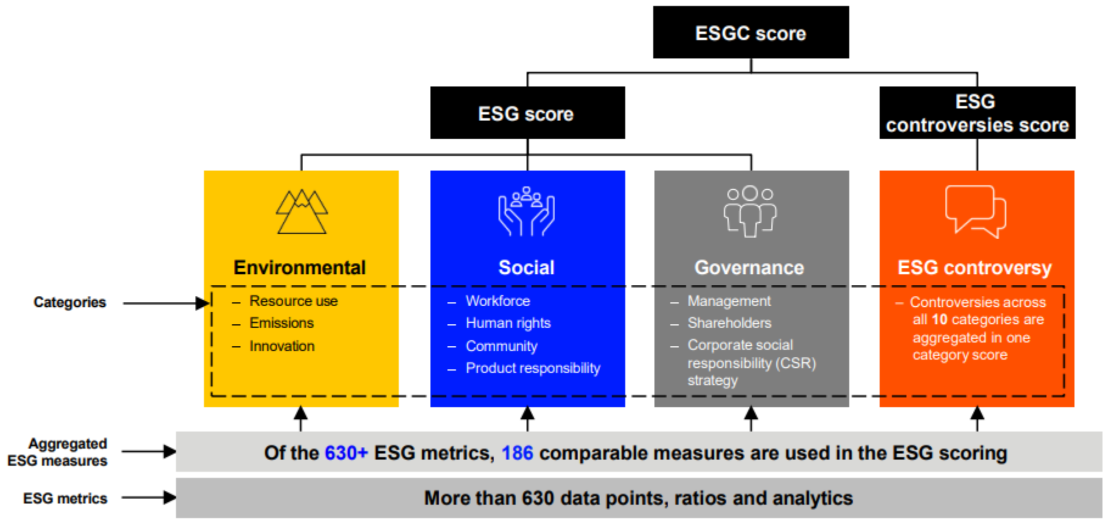

Figure 3: ESGC score composition overview

3.2 Fast ESG data construction

The fast ESG scores provided by MarketPsych, utilize advanced natural language processing (NLP) technologies to analyze real-time global media data. This approach captures dynamic Environmental, Social, and Governance (ESG) sentiments from news articles, blogs, tweets, and social media, offering a continuously updated snapshot that reflects the rapidly changing ESG landscape. MarketPsych's data sources include over 2,000 premium news sites such as Reuters, Bloomberg, and The Financial Times, as well as approximately 1,000 financial social media sites including Twitter, Facebook, LinkedIn, and influential blogs and forums like Medium and Reddit (MarketPsych, 2023; LSEG, 2023). This comprehensive coverage ensures that the sentiment scores reflect a wide array of opinions and information, providing investors with actionable insights.

MarketPsych's ESG scores start with the automated collection of vast volumes of text data from thousands of global content providers. This data is rich in ESG-related keywords, phrases, and contexts, and undergoes real-time sentiment analysis. Each piece of content is evaluated for relevance and sentiment polarity, with positive scores assigned to beneficial ESG practices and negative scores to detrimental events like industrial accidents or governance failures.

The scores are aggregated over various time windows, ranging from minutely to daily, depending on the score type in the Advanced or Core datasets. The Advanced dataset, updated minutely, tracks instantaneous public sentiment shifts across specific ESG indicators. In contrast, the Core dataset summarizes broader ESG categories into overall scores, updated daily to provide a comprehensive view of daily sentiment.

A distinctive feature of the scoring system is the inclusion of Buzz and Controversy scores. Buzz scores measure the volume of ESG-related media attention an entity receives, indicating the prominence of ESG issues. Controversy scores emphasize negative mentions and incidents, highlighting immediate ESG risks that may influence stakeholder perceptions or financial assessments.

These scores are dynamically updated as new data is processed, ensuring that they mirror the latest media narratives and public opinions. This real-time functionality is crucial in financial markets or industries where ESG perceptions shift rapidly, requiring stakeholders to swiftly adapt to new information.

3.3 Slow and Fast ESG data preparation

To analyse the differences between fast and slow ESG scores within the S&P 500, I adopted a methodology that aggregates fast scores into an annual format, matching the yearly cycle of slow scores (Refinitiv, 2023). This aggregation process involves a weighted average, where each daily score is adjusted according to its 'buzz' - a metric quantifying the volume and intensity of media coverage. This approach ensures that days receiving more media attention have a correspondingly larger impact on the annual score, accurately reflecting their significant influence on investor and public perceptions.

Technically, this is achieved by calculating the sum of the products of daily ESG scores and their respective buzz values, divided by the sum of all buzz values over the year. This Buzz-weighted average formula is applied to each company within the dataset, grouped by both company name and the year of the score data. The formula can be summarized as:

$$Annual\ Weighted\ ESG = \frac{\sum Daily\ ESG \times Daily\ Buzz}{\sum(Daily\ Buzz)}$$

Where:

- $Daily\ ESG$ is the ESG score for a given day.
- $Daily\ Buzz$ is the media buzz value for that day.
- The summation extends over all days in the year for each company.

This annualized ESG score adopts the format of the slow scores, enabling straightforward comparisons. Through this methodology, each year's aggregated fast score encapsulates the total effect of a company's ESG-related media visibility and public sentiment, smoothing out the daily fluctuations. This aggregation not only facilitates the evaluation of overarching trends and shifts in ESG perceptions throughout the year but also synchronizes the fast data with the slow data, ensuring a consistent framework for analysis. This alignment provides deeper insights into the long-term influence of ESG factors on company performance.

4. Project activities in cycles

In this chapter, I will discuss the cycles and steps undertaken to achieve the goals of this project. I will outline the planning and analysis methods used to ensure the desired results that were achieved. Additionally, I will explain how each step led us down different paths, ultimately providing a comprehensive review of my achievements. This will offer a complete picture of the project's process and progress.

4.1 Cycle 1: Foundations and initial analysis

4.1.1 Constructing

In the initial phase of my project, I established the foundation by outlining the theoretical framework, focusing on a comparative analysis of different types of ESG scores and their potential applications. I chose this topic to delve deeper into the ESG realm and understand how ESG scores impact today's financial environment. This topic is also significant due to the substantial financial influence that ESG scores currently have on companies.

The objective of the project was to explore the relationship between high-frequency ESG scores and traditional ones, and to determine how these scores can be used within S&P 500 companies. These applications include predicting ESG scores, explaining events that affect companies' ESG pillars, and investigating potential greenwashing practices. This research was further driven by my professional environment in a risk management consultancy and my desire to meet our clients' needs regarding ESG scores.

4.1.2 Planning

For the planning phase of the project, I developed a structured proposal to evaluate two different types of ESG scores. The data I gathered were provided by two financial firms in London. The slow ESG scores came from the Refinitiv database, covering the years 2005-2024 for all S&P 500 companies on a yearly basis. In contrast, the fast ESG scores were sourced from the MarketPsych database, spanning the years 1998-2024 on a daily basis.

Additionally, I obtained data for the Environmental, Social, and Governance pillars separately from 2005-2022 for the slow scores and from 1998-2024 for the fast scores. MarketPsych also provided controversy scores data for each S&P 500 company. One challenge I encountered, was missing data in the early years (2005-2009) from the Refinitiv database. MarketPsych also faced delays in data delivery, taking an entire month due to the large volume, which slowed the retrieval process.

These issues were addressed by communicating with both companies' helpdesks multiple times. Refinitiv experts managed to retrieve additional data for me, significantly reducing the percentage of missing values. On the other hand, a colleague from my organization, who is an expert in the field, assisted in optimizing the data retrieval process in my coding, enabling faster data process. These steps ensured that the data I had were sufficient for analysis.

4.1.3 Taking action

Slow and fast ESG scores can explain why companies are performing well or poorly over time. However, these scores can sometimes differ significantly at specific points. This discrepancy arises because slow scores are self-reported and published by companies, whereas fast ESG scores reflect market sentiment.

To gain more insights into the patterns and trends of these different data sources, I began with an exploratory data analysis. Before starting the analysis, I performed data cleaning, which was minimal since both datasets from the firms were delivered with solid data and few missing values. After cleaning the data, I standardized the formats, as the fast ESG scores were in a daily format and the slow scores were in a yearly format. I resolved this by converting the fast ESG scores to a yearly format.

For the exploratory data analysis, I used python to create visualizations, such as time-series graphs, ESG scores and pillar scores comparisons, correlation comparisons, and sector-specific comparisons. These visualizations helped uncover patterns and establish baselines for predictive modelling and my greenwashing methodology.

4.1.4 Evaluating action

Upon analysing the exploratory data, I observed connections and correlations between the fast and slow ESG data. Specifically, within the S&P 500 index, there are companies with differing ESG patterns. This observation was crucial in understanding the relationships between what companies report about their ESG actions and how society and people perceive those actions.

Reflecting on these findings and consulting with my supervisor, I realized that I needed to adjust my initial objectives. Initially, I planned to include a trading strategy based on high-frequency ESG scores in this thesis. However, I concluded that such a strategy would not be accurate using only ESG data, and the time was limited to search for additional parameters for all S&P 500 companies.

4.2 Cycle 2: Development and refinement

4.2.1 Building my product

In the first part of Cycle 2, I began building an ARIMA predictive model for the data I had collected and analysed. Due to the nature of the data and the problem I aimed to solve, I determined that an ARIMA model was the most suitable option. Although I also trained models such as the XGB (extreme Gradient Boosting) Regressor and a panel regression model, the results were not as effective as those produced by the ARIMA model. The ARIMA model fit the data better, which is why I included only this model in the paper, describing it in as much detail as possible.

With the ARIMA model, I aimed to evaluate the accuracy and variability of the slow ESG scores over the last five years (2020-2024). I achieved this by training the ARIMA model on both slow and fast ESG data from 2005 to 2020, attempting to capture relationships and patterns between these two data sources. By training my model with both data sources and performing the necessary fine-tuning, I was able to predict how much the slow ESG scores could change during the specified years, based on the values of the fast ESG scores in the same period. The metric used to measure this prediction accuracy was the Mean Absolute Error (MAE).

Finally, at the end of this thesis, I conducted a basic comparison of the two data sources for specific companies that met certain correlation criteria. This comparison aimed to understand how the data differed over time at the ESG level and across all three pillars (Environmental, Social, and Governance). This analysis revealed critical moments when companies' ESG practices were not aligned with societal perceptions.

4.2.2 Updating my product

While refining my ARIMA model, I engaged in deep conversations with both my university supervisor and my work supervisor. They recommended several changes to enhance my model's accuracy and avoid overfitting. Based on their advice, I restructured my data, training the model on data from 2005-2014 for both data sources and performing fine-tuning (parameter optimization) on a validation dataset comprising data from 2015-2020.

After identifying the most optimized parameters for each company in my dataset, I applied the same methodology with these parameters. This time, I trained the model using data from 2005-2019 and predicted outcomes for 2020-2024. These adjustments, based on my supervisors' guidance, were crucial for optimizing the machine learning process in my methodology.

For the final chapter of my thesis, I faced challenges in finding a methodology to measure greenwashing with my data. After several meetings with an ESG specialist in my professional environment, she provided me with valuable papers that significantly helped refine my methodology. Following a presentation of my methodology to her, she assisted in updating my strategy and provided insights on how to interpret specific events in my graphs as potential greenwashing indicators.

4.2.3 Finishing my product

I continued improving the model to enhance accuracy and optimize the code for faster execution. However, I realized that to significantly enhance the model's performance, I would need data from more companies and additional years. Regarding the greenwashing indicator, after finalizing my methodology and findings, I found that although I could pinpoint where and in which pillar issues arose for each company over time, I could not conclusively prove instances of greenwashing. Therefore, my product provides strong evidence but not definitive proof. In the future, enriching this strategy with more parameters and data could enable a more robust proof of greenwashing practices.

5. Project findings

5.1 Exploratory data analysis

Having detailed the collection and construction methods of MarketPsych's and Refinitiv's ESG data, I now turn to exploring their similarities and differences. By evaluating both data types, I aim to determine if fast ESG scores can effectively complement slow ESG scores, providing actionable insights for the S&P 500 companies.

5.1.1 Data analysis and cleaning

In this section, I conduct a review of ESG scores for companies within the S&P 500 universe, covering the years 2005-2024 for MarketPsych's Environmental (E), Social (S), Governance (G), and combined ESG scores, and the same period for Refinitiv's ESG scores, with Refinitiv's E, S, and G scores available until 2022. This analysis provides a foundational understanding of the behavior of these scores within this focused subset of roughly 500 companies, laying the groundwork for a more detailed study in the chapters that follow.

	Refinitiv (Slow Scores)				MarketPsych (Fast Scores)			
	E Pillar	S Pillar	G Pillar	ESG score	E Pillar	S Pillar	G Pillar	ESG score
Observations	7.751	7.751	7.751	10.060	3.155.040	3.155.040	3.155.040	3.155.040
Missing Values	1.231	1.231	1.231	1.775	54.383	1.266	11.107	0
Missing Values(%)	13,70	13,70	13,70	17,64	1,72	0,040	0,35	0
Mean	46,46	56,16	57,20	52,60	66,43	72,02	70,40	69,52
Median	50,58	57,55	59,53	54,23	72	77	75	73
Std. Dev.	28,80	21,62	21,41	20,15	22,65	18,98	20,78	15,90
Minimum	0	0.26	0.61	0,59	1	1	1	2
Maximum	98,54	98,25	99,44	95,16	100	100	100	99

Table 1: Overview of data availability and statistical analysis for S&P500 Companies

Table 1 highlights a significant difference between fast and slow ESG scores, particularly in data consistency. The slow ESG scores show many missing values because companies only publish their scores annually, making it harder to obtain comprehensive data. This is especially noticeable for smaller-cap companies in the S&P 500, where gathering accurate ESG data seems more challenging. In contrast, the fast scores, updated daily, have far fewer missing values due to their reliance on sentiment analysis and NLP techniques. Additionally, the next plot reveals that most missing values in the slow data are from earlier years, with the percentage of missing values dropping significantly in more recent years.

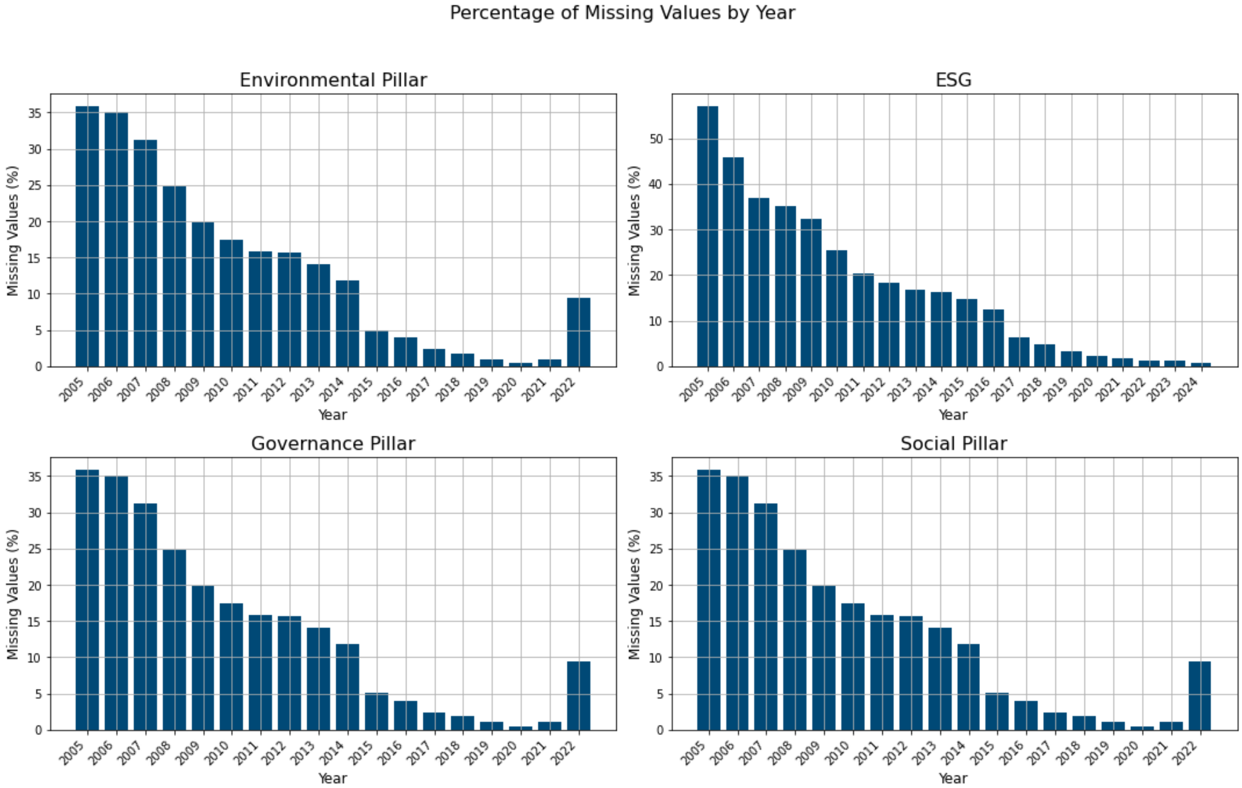

Figure 4: Percentage of missing values in Refinitiv dataset

For convenience and given the difficulty of consistently collecting slow ESG scores for all companies every year, I chose to use the S&P 500 companies as listed in January 2024. While companies frequently enter and exit the S&P 500 index, this approach provides a reasonable and representative dataset. I rely on the same company IDs from Refinitiv and MarketPsych to retrieve the data, ensuring consistency in the analysis. It is important to note that I do not focus on the current composition of the S&P 500, as my goal is not to analyse the index's present makeup but to have a representative sample of companies. Detailed information on these companies, as received from Refinitiv, can be found in Appendix 8.1 S&P 500 companies.

To tackle missing data in my study, I initially assessed the extent of missing values among S&P 500 companies. Companies with over 20% missing data were excluded to preserve data integrity and enhance model accuracy, reducing the dataset from 496 to 393 companies and decreasing the overall missing ESG data from 17.64% to 3.5%. This exclusion was informed by the concentration of missing values in smaller companies from earlier years. For the residual missing values, I applied interpolation, using subsequent year's data to fill gaps. For instance, missing ESG scores from 2005 were replaced with 2006 scores, ensuring all years from 2005 to 2024 were complete and minimizing bias.

5.1.2 Fast ESG scores

Before I move on to the comparison of slow and fast ESG scores, it is important to understand that the core of this thesis revolves around the fast ESG scores from MarketPsych. To provide an analytical representation, I have created key visualizations that display yearly average values for the Environmental, Social, and Governance pillars, along with ESG scores, combined ESG scores, and ESG controversies across the S&P 500.

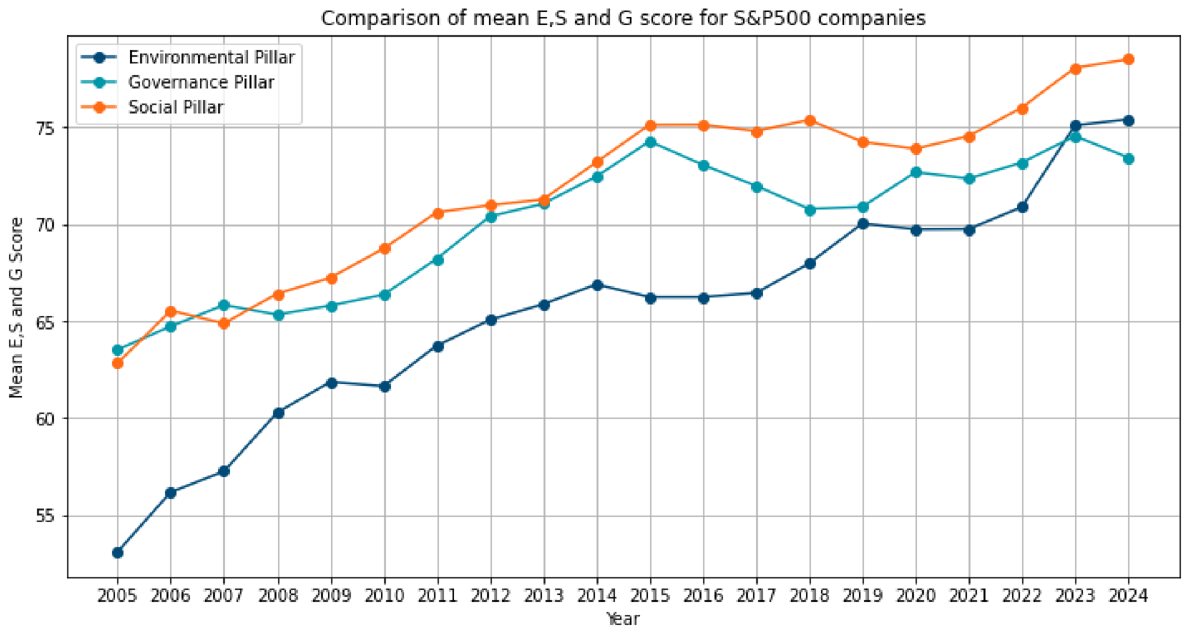

Figure 5: Mean Environmental, Governance and Social scores for S&P 500 companies

Figure 5 reveals that the Environmental (E) pillar consistently scores the lowest, followed by the Governance (G) pillar, while the Social (S) pillar outperforms the others. This trend suggests that companies might be facing more significant challenges or scrutiny in their environmental and governance practices, whereas social aspects are relatively better managed or perceived.

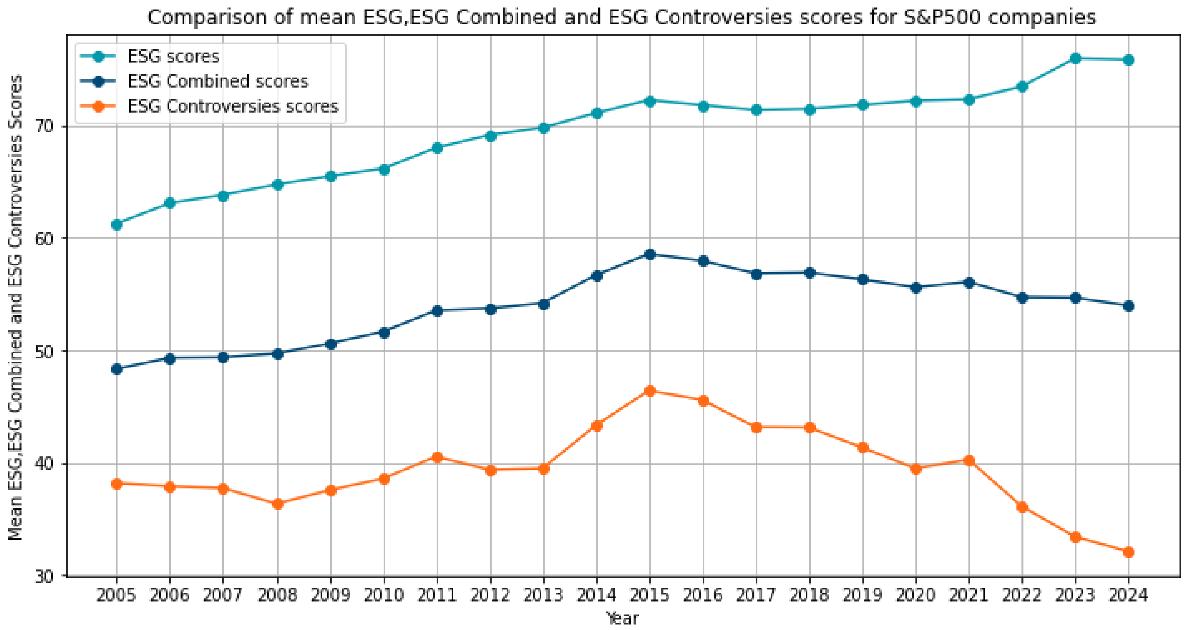

Figure 6: Mean ESG, ESG Combined and ESG Controversies scores for S&P 500 companies

Figure 6 presents the mean ESG scores, ESG combined scores, and ESG controversies scores for S&P 500 companies over the same period. The high ESG scores indicate strong overall performance, but the ESG combined scores show a noticeable reduction once ESG controversies are factored in. This reduction aligns with the methodology which states that if the ESG controversies score is higher than the ESG score, the combined score equals the ESG score. If the ESG score is higher, the combined score is an average of the ESG score and the controversies score. It is also noteworthy that after 2015, the ESG score shows minimal growth, while the controversy score decreases dramatically. This trend makes sense as controversies negatively impact public sentiment, making it more difficult for the overall ESG score to increase.

5.1.3 Data comparison

In the subsequent plot, I present the mean ESG values over time for the companies in the S&P500, covering both datasets.

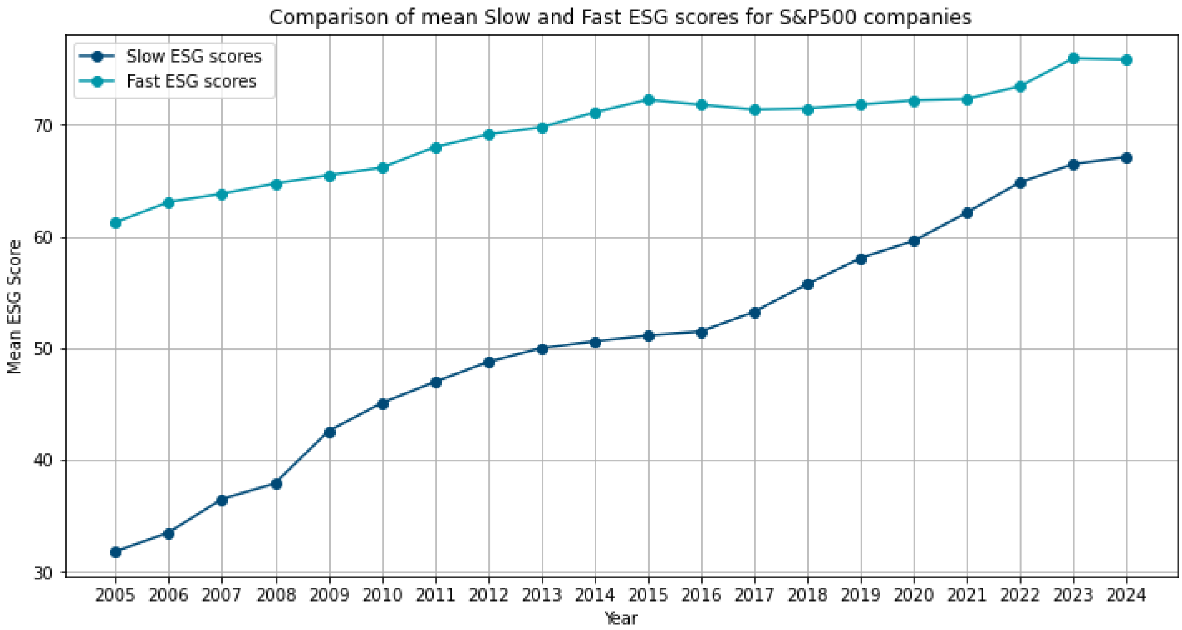

Figure 7: Mean ESG scores for S&P 500 companies

In Figure 7 we can see that both slow and fast ESG scores demonstrate an upward trend over the 20-year period, indicating a general improvement in ESG practices among the companies of S&P 500. Interestingly, while the fast scores started at a significantly higher level in 2005 compared to the slow scores, the slow scores exhibited greater growth over time. By 2024, the average difference between the slow and fast scores narrowed to just 9 points, down from a 30-point difference in 2005. It is also noteworthy that the fast scores remained relatively stable from 2015 to 2021, during which the slow scores continued to show substantial growth. This trend can be explained as companies tend to improve their self-reporting practices over time and becoming more proficient in communicating their ESG actions and goals. As companies have refined their reporting processes, the reliability and depth of slow ESG scores have increased, providing a more comprehensive reflection of their long-term ESG commitments.

5.1.4 Index correlation

Expanding on previous analyses of slow and fast ESG scores, I will now examine the correlation between these two datasets. To measure the relationship between the fast and slow ESG scores, I have calculated the Pearson correlation coefficient on an annual basis for each year starting from 2005. The Pearson correlation coefficient (r) is a statistical measure used to evaluate the strength and direction of the linear relationship between two variables. It ranges from -1 (indicating a perfect negative linear relationship), through 0 (no linear relationship), to +1 (a perfect positive linear relationship). The formula for r is as follows:

$$r = \frac{\sum(xi - \bar{x}) \times (yi - \bar{y})}{\sqrt{\sum(xi - \bar{x})^2 \times \sum(yi - \bar{y})^2}}$$

In this formula:

- xi and yi represent the annual aggregated fast and slow ESG scores for the S&P500 companies, respectively.
- $\bar{x}$ and $\bar{y}$ are the mean values of these scores for each respective year.

By calculating correlation coefficients annually and plotting them over time, we can see the evolving relationship between the two types of scores.

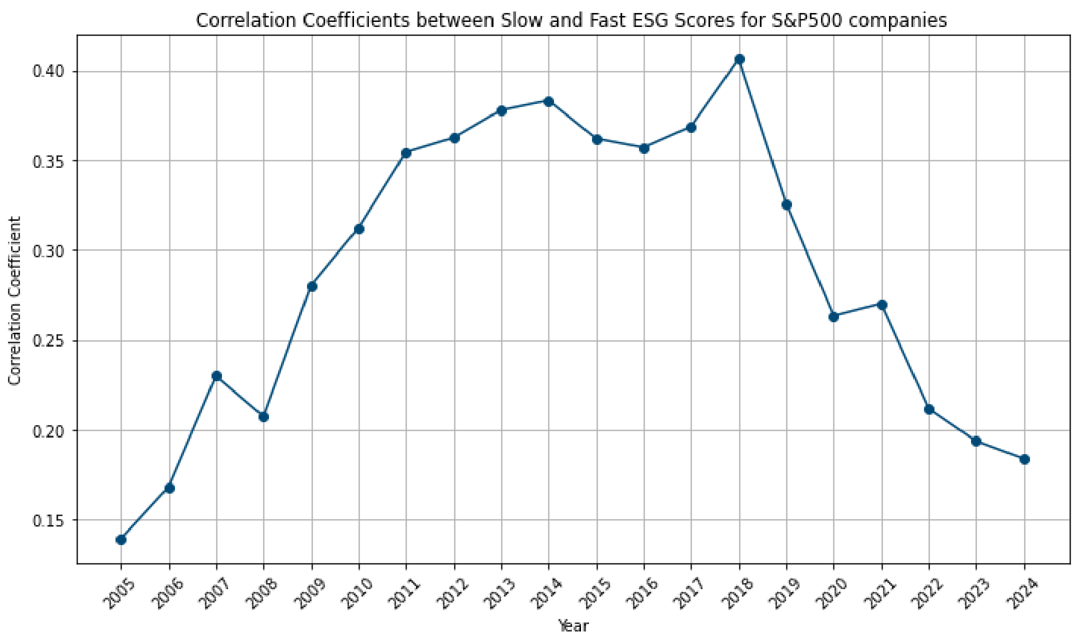

Figure 8: Correlation coefficients over time for slow and fast ESG data

Based on the Figure 8, the correlation between the ESG scores increases steadily from 2005 to around 2018, suggesting that during this period, fast scores became better aligned with slow scores, potentially indicating a convergence in the factors driving both types of scores. Notably, there are significant shifts in the correlation during specific periods which correspond to major global events. Between 2007 and 2008, the onset of the global financial crisis seems to affect the correlation, likely due to the immediate impact of financial instability on public sentiment captured by the fast scores, which is not as quickly reflected in the slow scores that are typically based on annual reporting cycles.

The sharp decline from 2018 to 2024 can be attributed to several factors. Increased regulatory scrutiny and new reporting standards, such as those introduced by the Task Force on Climate-related Financial Disclosures (Financial Stability Board, 2023), led companies to adjust their self-reported ESG data to meet heightened expectations, causing divergence from sentiment-driven scores. Additionally, market volatility and external shocks, particularly the COVID-19 pandemic, had immediate impacts on public sentiment that were more quickly reflected in fast scores than in the slower, annual reporting cycles. Moreover, the rise of social media and increased public awareness of ESG issues amplified immediate market reactions to corporate behaviors and events, often not reflected until later in self-reported data.

5.1.5 Companies correlation analysis

Following our assessment, I now will focus on a detailed company-level and yearly analysis. This deeper examination reveals how a company's real-time ESG performance (fast scores) aligns with its more systematically reported metrics (slow scores) over time.

To categorize the strength of the correlations, I set specific thresholds: a high correlation is defined as a Pearson correlation coefficient greater than 0,6; a moderate correlation ranges from 0,2 to 0,6; a low correlation spans from 0,0 to 0,2; and a negative correlation includes any values below 0,0.

Within these categories, I found that 121 companies exhibit a high correlation, 186 companies show a moderate correlation, while 70 companies are categorized under weak correlation. Finally, 63 companies display a negative correlation.

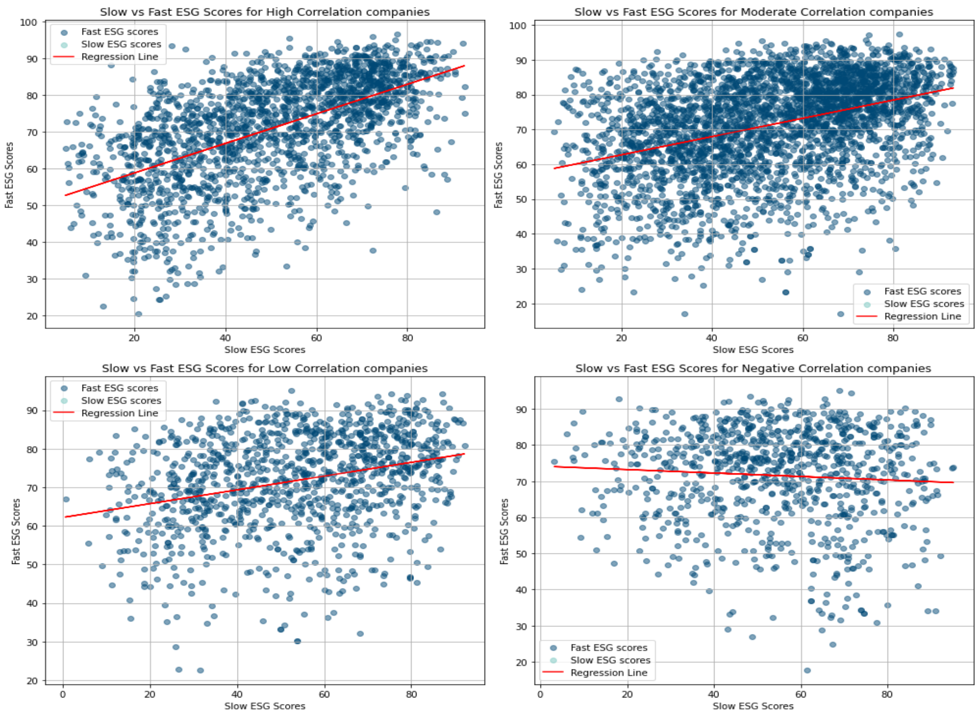

Figure 9: Slow vs. fast ESG score correlations

Figure 9 display the correlation between fast and slow ESG scores for different companies, highlighting varying levels of alignment. The 'High Correlation' plot shows a dense clustering along an upward trend, indicating that both scores increase together, reflecting consistent public perception and formal reporting. The 'Moderate Correlation' plot, while more dispersed, still shows a general upward trend, suggesting moderate alignment with occasional divergences due to reporting delays or differing aspects of ESG performance. The 'Low Correlation' plot features broadly dispersed data with no clear trend, indicating inconsistencies between public perception and formal reports. The 'Negative Correlation' plot suggests an inverse relationship, where fast scores increase as slow scores decrease, highlighting significant discrepancies potentially due to rapid shifts in public sentiment or unacknowledged issues in systematic reporting.

5.1.5.1 Sector-specific correlation characteristics

Continuing the analysis of the four correlation categories of S&P 500 companies, it is interesting to examine the companies sector within each correlation category. In the following table, I utilized the "Industry Name" from the fast ESG data to identify the number of distinct industries represented in each category and to highlight which industries are prominent within each category. A detailed analytical table is available in Appendix 8.5 Sectors and subsectors, where the reader can see the division of subsectors into the 11 sectors of the S&P 500.

Industry Category	Type of correlation	High	Moderate	Low	Negative
Communication Services		4	4	5	7
Consumer Discretionary		14	24	11	9
Consumer Staples		13	16	3	5
Energy		7	9	6	2
Financials		9	31	12	8
Healthcare		21	18	10	9
Industrials		22	39	11	11
Information Technology		21	35	12	11
Materials		4	9	5	4
Real Estate		10	10	2	3
Utilities		12	12	2	2
Unknown		6	6	3	2

Table 2: Industry Categories Overview by Correlation Group

The **Industrials** sector stands out with a high number of companies exhibiting positive correlations. Specifically, this sector has the highest number of companies with high and moderate correlations (22 and 39 respectively). This suggests that companies in the Industrials sector tend to have consistent ESG performance. One possible reason for this could be the nature of industrial companies, which often have more structured and regulated reporting processes due to the environmental and safety impacts inherent in their operations.

In contrast, the **Communication Services** sector shows a significant number of companies with negative correlations between their fast and slow ESG scores. With 7 companies showing negative correlations, it indicates a disparity between what companies report about their ESG practices and how they are perceived externally. This discrepancy could be due to the rapidly changing landscape of communication technologies and services, where public sentiment can be highly volatile and influenced by recent events or media coverage.

Consumer Discretionary and **Consumer Staples** sectors show substantial positive correlations, with Consumer Discretionary having 14 high and 24 moderate correlations, and Consumer Staples having 13 high and 16 moderate correlations. These sectors provide essential goods and services, which might lead to more consistent ESG reporting and perceptions. Companies in these sectors are likely to be scrutinized regularly, ensuring alignment between self-reported scores and public sentiment.

The **Energy** sector shows a mixed correlation pattern, with a notable number of companies in both high and moderate categories (7 and 9 respectively), but also a fair share in the low and negative categories. This variation might reflect the sector's ongoing transition towards sustainable energy practices, where newer, greener companies might show better alignment in their ESG scores compared to traditional oil and gas companies.

The **Financials** sector has a significant number of companies with moderate correlations (31), suggesting that while there is some alignment between reported and perceived ESG performance, there might still be discrepancies due to the complex nature of financial services and the variety of business models within this sector.

Healthcare also shows strong positive correlations with 21 high and 18 moderate correlations. This consistency might stem from the critical and highly regulated nature of health care services, ensuring that ESG practices are both rigorously reported and closely monitored by external sources.

Information Technology is another sector with strong positive correlations (21 high and 35 moderate), which could be due to the rapid innovation and transparency often associated with technology companies, as well as their significant impact on both societal and environmental factors.

Materials, **Real Estate**, and **Utilities** sectors have relatively fewer companies but still show a mix of correlations. Materials and Real Estate, with their tangible and highly regulated operations, tend to have clearer ESG reporting, while Utilities, being essential service providers, are often under stringent regulatory scrutiny, leading to more consistent ESG performance and perception alignment.

5.1.5.2 Social media attention and controversies

In addition to examining the industry categories and their distribution across correlation strengths, I also analysed the media attention (buzz) each correlation category received. Buzz serves as a proxy for media attention, comprising the number of ESG-relevant references to a given company in the media.

Type of correlation	Number of Companies	Buzz
High	121	3.124.965.896
Moderate	186	12.208.601.014
Low	70	6.367.579.180
Negative	63	14.024.474.080

Table 3: Media attention sum by correlation category

Table 3 reveals an interesting pattern. Companies exhibiting a negative correlation between their fast and slow ESG scores receive significantly higher media attention compared to other categories. This heightened attention is often due to controversies or significant events that swiftly alter public opinion. Initially, these companies may have high fast scores driven by immediate positive sentiment or hype. Over time, as systematic reporting catches up, their slow scores increase while fast scores decline due to waning public interest or emerging negative issues.

Conversely, companies with high correlation between their fast and slow ESG scores receive the least media attention. With almost one-fifth as many media mentions compared to the negative category, these companies tend to operate under the radar. Their consistent performance in both real-time sentiment and systematic reporting indicates stability and reliability in their ESG practices, which may not attract as much media attention. This relative lack of attention suggests that these companies successfully align their reported ESG initiatives with public perception, leading to fewer controversies or dramatic shifts in sentiment.

The media attention results raises the question: what is happening with the ESG scores and the controversies in each correlation category? In the following table, I used all available data from the fast ESG dataset to evaluate how controversies, category scores and the ESG pillars behave within each of the four correlation categories of the S&P 500 companies. This was done by calculating the average values of each category for all companies within each correlation group.

Category	Type of correlation	High	Moderate	Low	Negative
ESG		67,87	69,56	70,17	71,17
ESG Combined		55,13	54,82	52,94	53,26
ESG Controversies		44,58	41,18	36,55	35,80
Environmental Pillar		63,74	66,22	67,59	68,31
Governance Pillar		69,85	70,14	71,06	71,79
Social Pillar		70,50	72,39	72,39	73,62
CSR Strategy		63,62	61,60	64,83	65,49
Community		79,68	79,87	78,42	78,74
Emissions		62,04	63,86	65,61	66,12
Environmental Innovation		69,06	72,09	73,10	73,45
Human Rights		64,15	65,66	67,14	68,90
Management		74,14	74,77	75,45	75,42
Product Responsibility		68,94	71,48	69,18	69,04
Resource Use		64,15	65,83	66,65	67,18
Shareholders		60,51	60,28	60,78	63,31
Workforce		71,91	73,43	75,12	77,86

Table 4: Categories scores sum by Correlation Category

In Table 4, we can observe that although all the scores across the categories are quite similar, negatively correlated companies clearly perform better in their categories scores. The overall ESG score is the highest and the other categories representing each of the E, S, and G pillars generally show better scores. (For a detailed explanation of what these categories represent, please refer to the Appendix 8.2 Categories table).

Finally, looking at the data and results, we can see that the first three correlation categories (High, Moderate, and Low) show similar patterns, which are present in Appendix 8.3 Correlation categories between Fast and Slow ESG scores. In these categories, both fast and slow ESG scores generally rise together through time, although at different rates. However, the negative correlation category behaves differently. In this category, fast ESG scores decrease over time while slow ESG scores increase. This difference is especially interesting when we see how close these scores get in 2024.

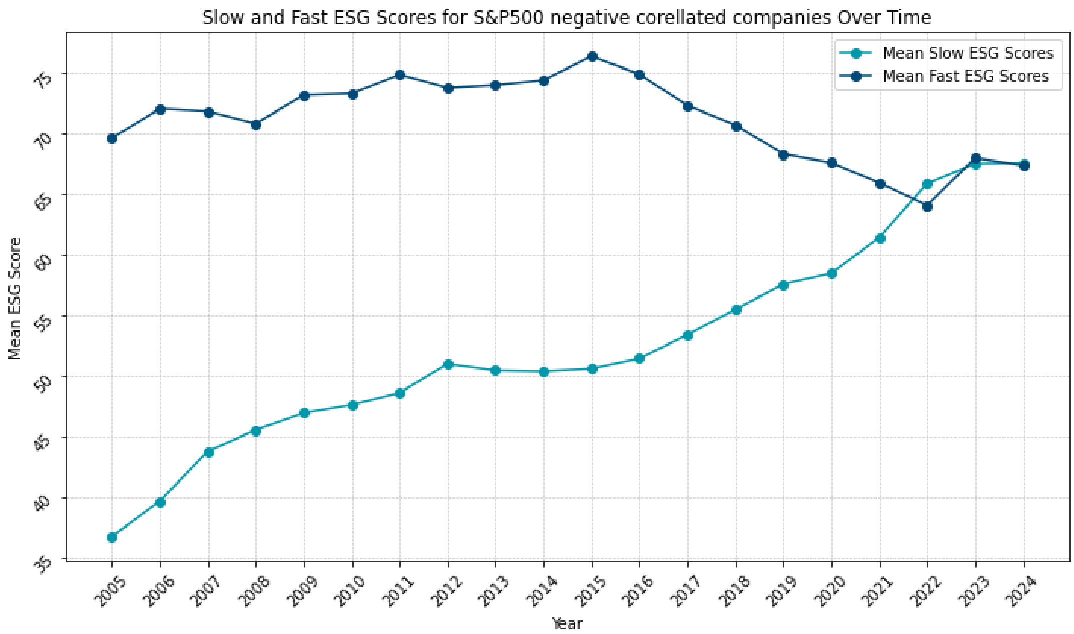

Figure 10: Negative correlation between ESG scores

Based on Figure 10, slow ESG scores steadily rise and stabilize near 70 by 2024, while fast scores remain stable with a slight increase until 2015, then decline sharply, aligning with slow scores by 2024. This trend indicates a disconnect between public sentiment and self-reported ESG performance starting in 2015. By 2024, the convergence of these scores around 70 reflects improved ESG reporting accuracy, stabilization of public perception, and better alignment with actual performance. Additionally, advances in AI may have enhanced sentiment data collection, contributing to consistent findings from both data sources by 2024. The decline in fast ESG scores after 2015 highlights the significant impact of public perception and media on company reputations.

5.1.5.3 Correlation characteristics by company market cap

In this section, I analysed the correlation of S&P 500 companies based on their market capitalization. Having previously examined the companies' characteristics, sector affiliations, and media attention, I now turn to how these companies' correlations manifest when grouped by market cap.

Due to the lack of market cap data from Refinitiv and MarketPsych, I manually searched for the market cap values of all companies in my dataset as of March 2024. The market cap of these companies ranges from $2,8 trillion (Apple Inc) to $2,13 billion (O'Reilly Automotive Inc), with a total market cap of approximately $4,1 Trillion USD.

I maintain the same correlation thresholds as in my previous analysis, categorizing correlations as high, moderate, low, or negative. Splitting the total market cap into three equal buckets, starting with the highest market cap companies, resulted in the following outcomes:

Market Cap Range	Correlation	Strong	Moderate	Weak	Negative	Number of Companies	Bucket Worth(B)
High (2,8T-468B)		2	5	2	3	12	13.703,14
Medium (468B-98B)		10	19	12	16	57	13.703,14
Low (98B-2,13B)		109	161	55	44	369	13.703,14

Table 5: Market cap and correlation results revised

Large companies (Market Cap: $2,8T to $468B): Out of 12 companies, only 2 show strong alignment between reported data and public sentiment, while 3 actually report better scores as sentiment declines. It seems that even with more resources, large companies sometimes struggle to keep their public image in sync with internal reports.

Medium-sized companies (Market Cap: $468B to $98B): This group is all over the place. With 57 companies it total, we see a real mix—10 are in strong agreement, 19 somewhat aligned, 12 barely aligned, and 16 moving in opposite directions. It looks like these companies are still figuring out how to effectively integrate and report on ESG.

Smaller companies (Market Cap: $98B to $2,13B): Smaller companies are generally doing a better job at aligning their reports with public sentiment, with 109 companies showing strong agreement. However, 44 companies show the opposite trend, which might reflect the challenges they face in maintaining consistent ESG practices due to limited resources.

5.2 Predictive model

In the domain of Environmental, Social, and Governance (ESG) metrics, stakeholders such as investors and policymakers rely heavily on the stability and accuracy of these scores for informed decision-making. Refinitiv notes that the slow ESG scores over the last five fiscal years, in our case from 2020 to 2024, are non-definitive and subject to change based on new information and corporate disclosures. This uncertainty highlights the need for more reliable predictive tools.

In this section, I will investigate whether fast ESG scores can predict changes in the slow ESG scores. Employing ARIMA modelling, I can evaluate the ability of fast scores to forecast future revisions within a five-year window by capturing patterns between the two different data sources.

5.2.1 ARIMA framework

With the dataset ready for modelling, each row corresponded to the ESG data of a unique company from 2005 to 2024. This required determining the optimal ARIMA parameters (p,d,q) for each company individually to ensure that the model trained on one company's data did not influence the forecasts for another.

In this phase, I will describe the three main parameters (p,d,q) in our model and how they are calculated.

- p (AutoRegressive order): In our context, p represents the number of lagged slow ESG scores (S) included in the model. This parameter captures the linear relationship between the current slow ESG score and its historical values. Mathematically, it is represented as:

$$S_t = c + \varphi_1 S_{t-1} + \varphi_2 S_{t-2} + .. + \varphi_\rho S_{t-\rho} + \varepsilon_t$$

where S_t the current slow ESG score, $S_{t-1}, S_{t-2}, .., S_{t-\rho}$ are the lagged slow ESG scores, $\varphi_1, \varphi_2, .., \varphi_\rho$ are the autoregressive coefficients, c is a constant term, and ε_t is the error term. The selection of p is achieved through an automated grid search process, testing various combinations and evaluating them using the Akaike Information Criterion (AIC) for the best fit.

- d (Integrated order): This parameter represents the number of times the raw observations of slow ESG scores need to be differenced to achieve stationarity. In simpler terms, it is the number of differences needed to remove trends or seasonality. Mathematically, differencing is represented as:

$$\tilde{S}_t = S_t - S_{t-1}$$

where $\tilde{S}_t$ the differenced slow ESG score. The parameter d is optimized automatically as part of the ARIMA modelling process, testing values from 0 to 2 and choosing the one that results in the best model fit.

- q (Moving Average order): In my model, q represents the number of lagged forecast errors included in the model. It captures the linear dependency between an observation and residual errors from a moving average model applied to lagged observations. Mathematically, it is represented as:

$$S_t = c + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + .. + \theta_q \varepsilon_{t-q} + \varepsilon_t$$

where $\varepsilon_{t-1}, \varepsilon_{t-2}, .., \varepsilon_{t-q}$ are the residual errors, $\theta_1, \theta_2, .., \theta_q$ are the moving average coefficients, and all other terms are as defined before. The selection of q is also part of the automated grid search process, with the best combination determined based on the AIC.

Training and validation data (Fine tuning)

1. Training data (2005-2015):
   - The ARIMA model is trained using pairs of slow ESG scores (S) and fast ESG scores (F) from 2005 to 2015. This means for each year within this range, the model uses the corresponding slow and fast ESG scores to capture patterns for each company separately.
   - Example: The model uses the data pairs $(2005\mathrm{S}, 2005\mathrm{F}), (2006\mathrm{S}, 2006\mathrm{F}), \ldots, (2015\mathrm{S}, 2015\mathrm{F})$ to understand the relationship between slow and fast scores for every company.
2. Validation data (2016-2020):
   - The model's ability to predict slow ESG scores is validated against the actual scores from 2016 to 2020. This validation phase involves predicting the slow ESG scores for each year in this range using the fast ESG scores of the same years.
   - Example: To predict the $2016\mathrm{S}$ score, the model uses the $2016\mathrm{F}$ score; for $2017\mathrm{S}$, it uses $2017\mathrm{F}$, and so on.

Additionally, I also trained and fine-tuned the model using data from 2010 to 2015 to compare the results with the 2005-2015 training period and to identify the best parameters. By evaluating both periods, I ensured the selection of optimal parameters that enhance the model's predictive performance.

After determining the optimal ARIMA parameters (p, d, q) for each, I applied these parameters to predict the non-definitive ESG scores for the years 2020-2024. This process involves using a rolling window ARIMA model, where the model is retrained each year with the latest available data. This ensures that the model uses the most recent information for each prediction.

Let D_t represent the training data available up to year t.

For each prediction year T from 2020 to 2024, we train the ARIMA model on D_{T-1}, which includes pairs $(S_{2005}, F_{2005}), (S_{2006}, F_{2006}), \dots, (S_{T-1}, F_{T-1})$.

The prediction for the slow ESG score $\hat{S}_T$ is made using the fast ESG score F_T:

$$\hat{S}_T = f(F_T)$$

where f represents the prediction function modeled by ARIMA.

Let's assume we have a company with corresponding slow and fast scores for the years 2005-2024. Here is how our ARIMA algorithm works:

1. Prediction for 2020:
   a. Training data: $(S_{2005}, F_{2005}), (S_{2006}, F_{2006}), \dots, (S_{2019}, F_{2019})$.
   b. Prediction: $\hat{S}_{2020} = f(F_{2020})$
2. Prediction for 2021:
   a. Training data: $(S_{2005}, F_{2005}), (S_{2006}, F_{2006}), \dots, (S_{2020}, F_{2020})$.
   b. Prediction: $\hat{S}_{2021} = f(F_{2021})$
3. Prediction for 2022:
   a. Training data: $(S_{2005}, F_{2005}), (S_{2006}, F_{2006}), \dots, (S_{2021}, F_{2021})$.
   b. Prediction: $\hat{S}_{2022} = f(F_{2022})$
4. Prediction for 2023:
   a. Training data: $(S_{2005}, F_{2005}), (S_{2006}, F_{2006}), \dots, (S_{2022}, F_{2022})$.
   b. Prediction: $\hat{S}_{2022} = f(F_{2022})$
5. Prediction for 2024:
   a. Training data: $(S_{2005}, F_{2005}), (S_{2006}, F_{2006}), \dots, (S_{2023}, F_{2023})$.
   b. Prediction: $\hat{S}_{2023} = f(F_{2023})$

In addition to use the 2005-2019 period for training, I also used the periods 2010-2019 and 2015-2019 to make predictions for the years 2020-2024. This allowed me to compare the results and ensure the robustness of the model.

To summarize the effectiveness of the ARIMA model in predicting the ESG scores for the last five years (2020-2024), I consider both the quantitative metrics and the visual results displayed in the plots.

1. For my ARIMA model, the MAE measures the average magnitude of errors between the predicted ESG scores and the actual observed scores. It is calculated as the average of the absolute differences between each pair of predicted and actual values. This metric helps in understanding the overall accuracy of my model without considering the direction of the errors.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

2. The MSE calculates the average of the squared differences between predicted and actual ESG scores. By squaring the errors, the MSE penalizes larger errors more than smaller ones. This metric is particularly useful for identifying models with significant outliers or large deviations.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

3. The RMSE for ARIMA model is the square root of the MSE, providing an estimate of the standard deviation of the model's prediction errors. It gives a sense of how much prediction errors deviate from the actual values on average. This metric makes it easier to interpret the prediction error in the same units as the original data.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

4. MAPE in my ARIMA model measures the accuracy of the forecasts as a percentage, providing a clear perspective on the relative scale of errors. It is useful for understanding how significant the prediction errors are in relation to the actual values. This metric is particularly valuable when comparing the accuracy across different scales or datasets.

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right|$$

5. The SEM quantifies the precision of the sample mean estimate of the population mean, indicating the spread of the sample mean's distribution in my ARIMA model predictions. This metric helps in understanding the reliability of the sample mean of the prediction errors.

$$SEM = \frac{\sigma}{\sqrt{n}}$$

where σ is the standard deviation of the sample of prediction errors.

6. The adjusted R-squared measures the proportion of the variance in the ESG scores that is predictable from the independent variables (fast ESG scores), adjusted for the number of

predictors in my ARIMA model. This metric adjusts the R-squared value based on the number of predictors, providing a more accurate measure of the model's explanatory power.

$$Adjusted\ R^2 = 1 - (1 - R^2)\frac{n-1}{n-p-1}$$

where n is the number of observations and p is the number of predictors.

5.2.2 ARIMA results

After training ARIMA model and performing cross-validation on the validation test set, I obtained the following results:

Metric	Training set years	2005-2015	2010-2015
Mean Absolute Error (MAE):		131,99	35255,76
Mean Squared Error (MSE):		497066,48	296860083353,20
Root Mean Squared Error (RMSE):		705,02	544848,67
Mean Absolute Percentage Error (MAPE)		242,87	68460,58

Table 6: Parameters results

Table 6 indicate that the training set from 2005-2015 yields better performance metrics compared to the training set from 2010-2015. Specifically, the 2005-2015 training set demonstrates significantly lower MAE, MSE, RMSE, and MAPE values, suggesting a more accurate and reliable model. Therefore, I will use the parameters derived from the 2005-2015 training set for our predictions.

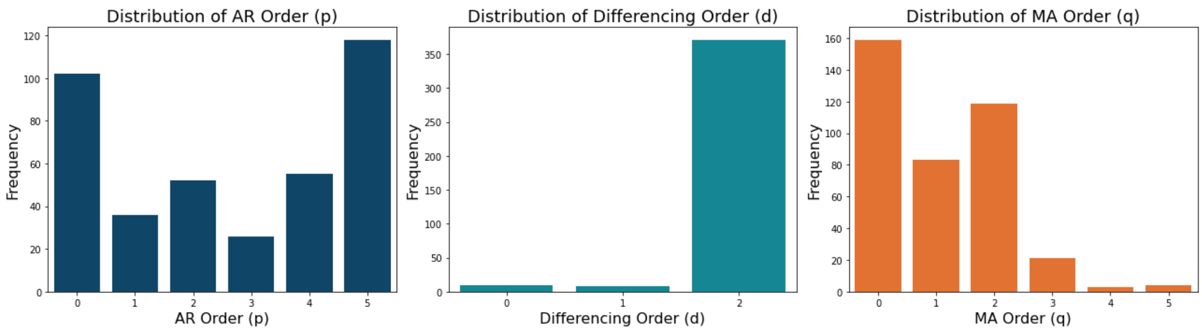

Figure 11: Distribution of AR, MA, differencing orders

- AR Order (p): The distribution of the AR (autoregressive) order shows a varied preference across different companies. The most frequent AR orders are 0 and 5, indicating that in many cases, either no lag or five lagged observations significantly improved the model fit.

- Differencing Order (d): The differencing order is predominantly 2 for most companies. This suggests that taking the second difference of the time series was necessary to achieve stationarity in the majority of the datasets.
- MA Order (q): The distribution of the MA (moving average) order shows a strong preference for values 0 and 1. This indicates that either no lagged forecast error term or a single lagged error term was sufficient for the majority of the models.

The results from Table 7 indicate that the model trained on the data from 2005-2019 significantly outperforms the models trained on shorter time periods (2010-2019 and 2015-2019). The metrics show that incorporating a longer historical dataset leads to better model performance. Specifically, the training set from 2005-20219 achieves lower errors and a higher adjusted R-squared, demonstrating more accurate and reliable predictions.

Metric	**Training set years**	**2005-2019**	**2010-2019**	**2015-2019**
Mean Absolute Error (MAE):		6,94	11,42	271,83
Mean Squared Error (MSE):		83,91	8496,92	74557554,82
Root Mean Squared Error (RMSE):		9,16	92,17	8634,67
Mean Absolute Percentage Error (MAPE)		11,31%	18,23%	377,60%
Standard Error of the Mean (SEM)		0,21	2,09	196,62
Adjusted R-Squared:		0,56	-43,76	-395147,55

Table 7: Overview of ARIMA metrics results

The Mean Absolute Error (MAE) of 6.94 indicates that, on average, the model's predictions deviate from the actual values by approximately 6.94 points, suggesting a reasonably good fit given the data range. The Root Mean Squared Error (RMSE) of 9.16 further confirms the model's accuracy, as it emphasizes larger errors by squaring them, providing a slightly more sensitive measure than MAE. The Mean Absolute Percentage Error (MAPE) of 11.31% reveals that the model’s predictions typically fall within 11.31% of the actual values, offering a clear perspective on the relative error scale. The Standard Error of the Mean (SEM) of 0.21 indicates a relatively low spread of the sample mean's estimation, reinforcing the model's reliability. Additionally, the adjusted R-squared value of 0.56 suggests that approximately 56% of the variability in the ESG scores is explained by the model, which is commendable given the complexity of ESG score dynamics.

Visual Evaluation

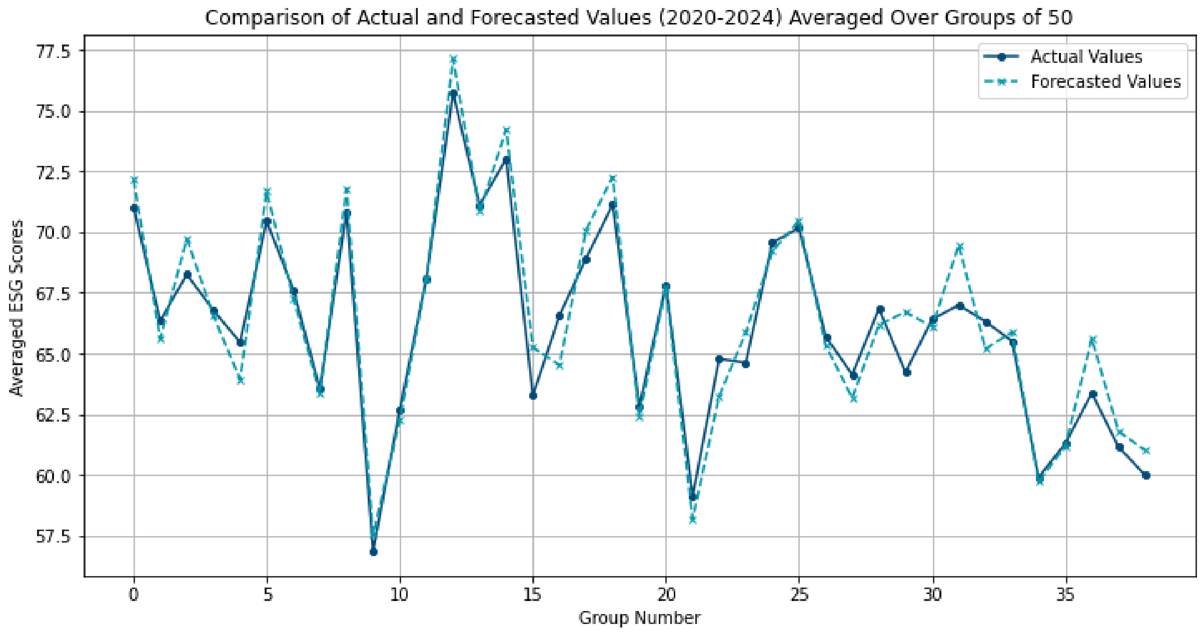

Figure 12: Actual vs. forecasted ESG scores for S&P 500 companies (2020-2024)

Figure 12 shows the comparison between the actual ESG scores and the forecasted values by the ARIMA model over the period 2020-2024, averaged over groups of 50. The close alignment of the lines indicates that the model's predictions are generally accurate. The actual and forecasted values follow a similar trend, demonstrating the model's ability to capture the underlying patterns in the ESG scores effectively.

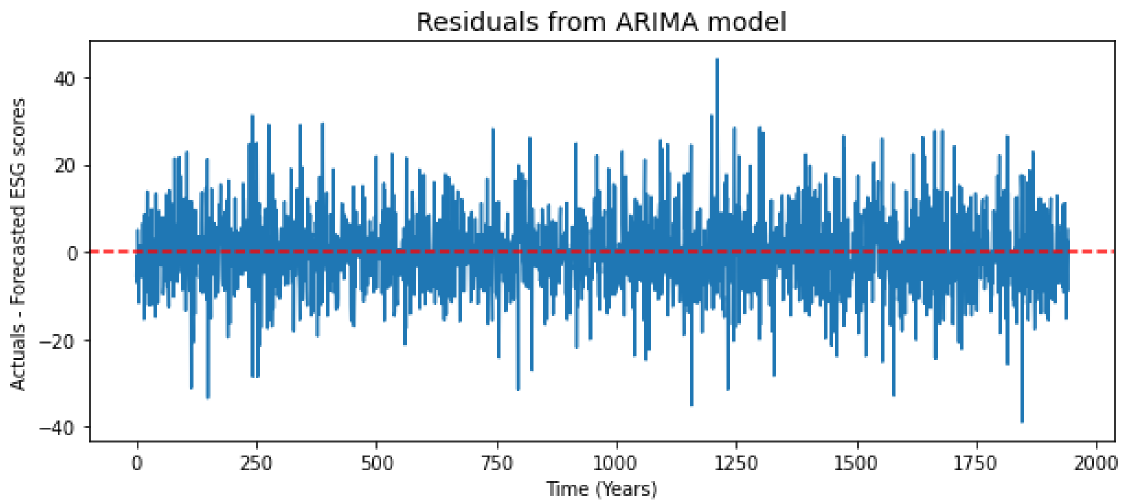

Figure 13: Residuals from ARIMA model

Figure 13 displays the differences between the actual ESG scores and the forecasted values over time. The residuals are centered around zero, which suggests that there is no systematic bias in the model's predictions. However, there are periods where the residuals are more dispersed, indicating potential areas where the model could be improved. Overall, the distribution of residuals indicates that the model performs well, but there is room for further refinement.

Conclusion

The ARIMA model demonstrates a solid capability to predict the non-definitive ESG scores of the last five years, which are subject to adjustments and restatements. The statistical measures indicate a good fit, while the plots visually corroborate that the predicted values align well with the actual data. The results show that fast ESG data have predictive power over the slow ones. This capability is particularly useful for financial analysis and investment decisions, as understanding potential future adjustments to ESG scores can provide a competitive edge. However, it is important to note that although the model performs quite well with the given data, there is still an average error of 6.94 units in the slow ESG scores over the last five years. This deviation is significant considering we are dealing with ESG scores of S&P 500 companies. Also this error suggests that investors should be cautious when using ESG data from Refinitiv for specific companies, as these scores can fluctuate by an average of 6.94 units over time. Furthermore, it is worth mentioning that with additional yearly data and data from more companies, there is a high likelihood that the model's accuracy will improve, as evidenced by our cross-validation results.

5.3 Greenwashing indicator

Greenwashing is a deceptive practice where companies exaggerate or misrepresent their environmental actions to appear more sustainable than they actually are. This practice can mislead stakeholders, including investors and regulators who rely on accurate representations of a company's environmental responsibility. A Greenwashing Indicator serves as a critical tool to identify and quantify discrepancies between a company's publicly promoted environmental image and its actual environmental practices.

In this section, I will outline the methodology utilized by Covalence SA in their Greenwashing Indicator (Covalence SA, 2020) and explore how I can adapt this methodology to Refinitiv and MarketPsych ESG data.

5.3.1 Literature review and Covalence's methodology

Covalence is a company that specializes in evaluating corporate practices to uncover discrepancies between companies' proclaimed and actual environmental, social, and governance (ESG) behaviors. Covalence's methodology employs a dual-data approach that integrates both quantitative and qualitative elements. Quantitative data are primarily derived from disclosures that companies publish annually. These disclosures include ESG and SDG[3] indicators, such as health and safety policies in the supply chain, the percentage of women on the board, and total water use per million in revenue, sourced from external providers and the companies themselves. Annual disclosures enter the indicator by contributing quantitative data that is used to calculate this disclosure score, which is then combined with a sentiment-based reputation score to form a comprehensive ESG rating. This rating helps in identifying discrepancies between a company's stated commitments and actual performance, thus informing the Greenwashing Risk Indicator. The qualitative component, on the other hand, comes from news-based sources. Covalence evaluates various news items from media, NGOs, trade unions, governments, and other stakeholders, assigning each a sentiment score based on its positive or negative portrayal of the company.

[3] SDG indicators: Metrics used to measure progress towards the United Nations Sustainable Development Goals.

Calculation of greenwashing Risk

Using this categorized news data, Covalence computes a reputation score for each category. The score is an aggregate of the news sentiments - positive news contributes to a higher score, suggesting favourable public perception, whereas negative news lowers the score, indicating controversies or poor performance.

The Greenwashing Risk Indicator itself is then derived by measuring the discrepancy between the forward-looking and backward-looking sentiments. A significant difference where the forward-looking sentiment exceeds the backward-looking sentiment might indicate greenwashing, as it suggests that the company's public commitments are not matched by tangible actions. Conversely, if the backward-looking sentiment is higher, it can suggest either green muting—where the company's actions are not sufficiently communicated or recognized—or that past negative perceptions are overshadowing current positive initiatives.

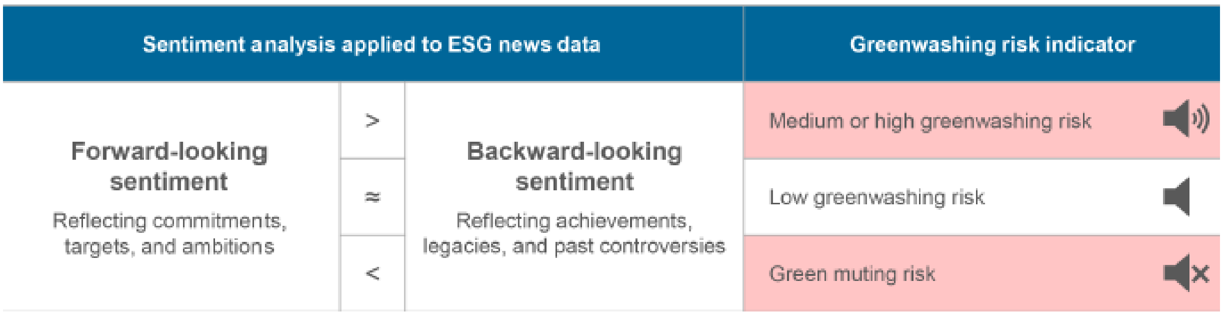

Figure 14: Covalence sentiment analysis and greenwashing risk

5.3.2 Fast and Slow ESG scores methodology

Building on the approach developed by Covalence, my methodology similarly examines the alignment or disparity between fast (MarketPsych) and slow (Refinitiv) scores. In this context, fast scores can be seen as a proxy for forward-looking sentiment reflecting public and media perceptions about a company's sustainability commitments. Slow scores, representing self-reported data, align with backward-looking sentiment that details a company's actual sustainability practices and achievements. By looking the discrepancies between what the public thinks about a company and what the company reports about itself, I can assess the authenticity of a company's ESG claims.

Detailed breakdown of my methodology

1. Low risk:
    - Definition: This scenario occurs when there is a high degree of similarity between Slow and Fast ESG scores.
    - Implication: Such alignment, suggests that the company's self-reported data accurately reflect public and media perceptions, indicating transparency and genuine compliance with ESG standards.

2. Sentiment misalignment:

Fast score higher than slow score:

- This situation indicates a potential overestimation of sustainability efforts in public or media narratives, not fully supported by the company's actual, reported practices. Company might be benefiting from an exaggeratedly positive public sentiment despite real issues, or that companies might intentionally understate their sustainability achievements as a strategic decision to under promise and overdeliver, or to delay drawing attention until certain outcomes are guaranteed.

3. Greenwashing:

Slow score higher than Fast score:

- This suggests possible overstating of sustainability efforts in official reports. If this pattern persists over time without corresponding support from public sentiment, it indicates potential greenwashing.

In the following section, I will apply this methodology to examine greenwashing risks within four S&P 500 companies, each representing different categories of correlation between their Slow and Fast ESG scores as defined in section 5.1.5. I chose one company from each correlation category to capture the varying patterns between the two data scores and to determine the extent to which these patterns are extreme or not. These companies are:

Company	**Pearson correlation**
Visa Inc.	0,86 (High)
NVIDIA	0,53 (Moderate)
JP Morgan Chase & Co	0,13 (Low)
Meta	-0,77 (Negative)

Table 8: Correlation results

5.3.2.1 Visa Inc.

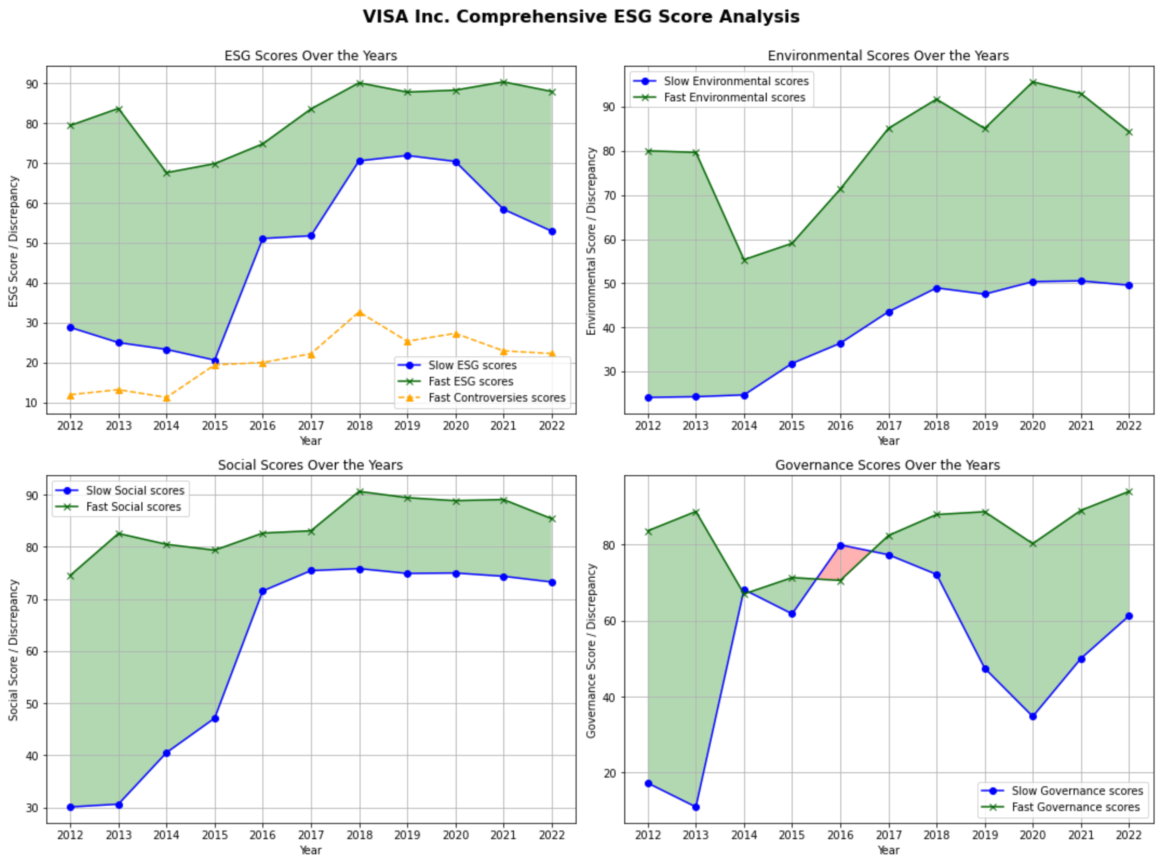

Figure 15: Comprehensive ESG score analysis for Visa Inc

In examining Visa Inc.'s ESG scores from 2012 to 2022, significant discrepancies are evident across all components (Environmental, Social, Governance), with media sentiment-based scores consistently exceeding self-reported scores, especially in later years. This pattern points to a medium to high risk of greenwashing. The persistent, substantial gaps suggest that media narratives may portray the company's sustainability efforts as more robust than indicated by their reports. This trend underscores potential risks of greenwashing, where actual sustainability actions are not align with public perceptions or presentations, highlighting the need for Visa to reassess and potentially adjust its reporting and communication strategies to accurately reflect its ESG initiatives.

5.3.2.2 NVIDIA

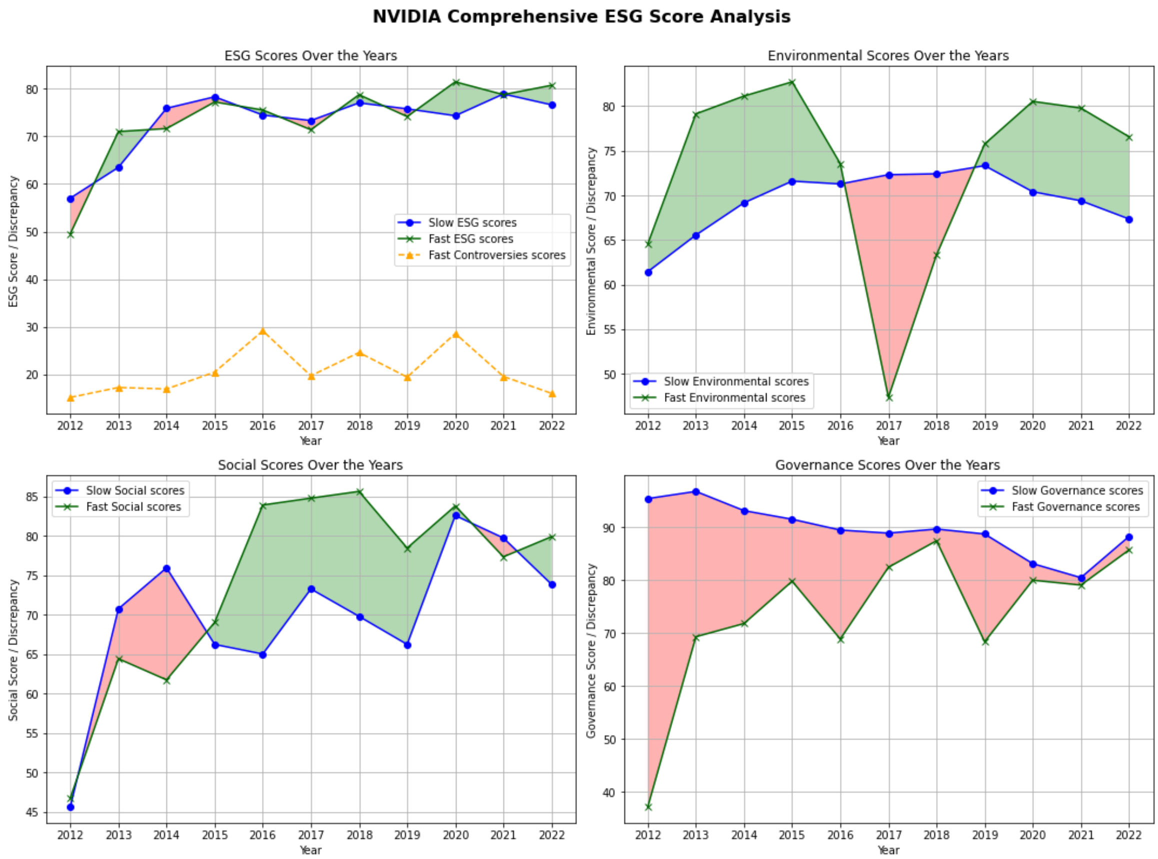

Figure 16: Comprehensive ESG score analysis for NVIDIA

From 2012 to 2022, NVIDIA's ESG scores reveal complex discrepancies across its Environmental, Social, Governance, and overall components scores. The overall ESG alignment is generally good, though some years show minor discrepancies, indicating low greenwashing risk. The Environmental scores fluctuate significantly, with Fast scores occasionally surpassing Slow scores, and vice versa, suggesting variable public perceptions or reporting adjustments that may lead to intermittent greenwashing or sentiment misalignment.

A significant event in 2017 highlighted these issues. A marked drop in fast Environmental scores coincided with increased scrutiny of NVIDIA's energy use during the crypto mining surge, when many of NVIDIA's GPUs were used for crypto mining, substantially increasing energy consumption and environmental impact. This period lacked clear disclosure about the effects of crypto mining on their operations, negatively impacting public sentiment and fast ESG scores (U.S. Securities and Exchange Commission, 2022). Despite NVIDIA's investments in climate solutions like the Earth-2 supercomputer for climate modelling, the immediate environmental impacts from their expanding operations and energy use likely heightened concerns (NVIDIA, 2022). The Social (S) scores for NVIDIA also show variability, reflecting dynamic public sentiment that sometimes diverges notably from reported data, indicating a medium risk of greenwashing. The Governance (G) scores the early years display substantial discrepancies, with Slow scores considerably higher than Fast scores; these gaps narrow significantly over time, suggesting a decreasing risk of green washing. This complex pattern across categories underscores the need for NVIDIA to improve transparency and communication to better align reported data with public perceptions and effectively address the potential risks of greenwashing and sentiment misalignment.

5.3.2.3 JP Morgan Chase & Co.

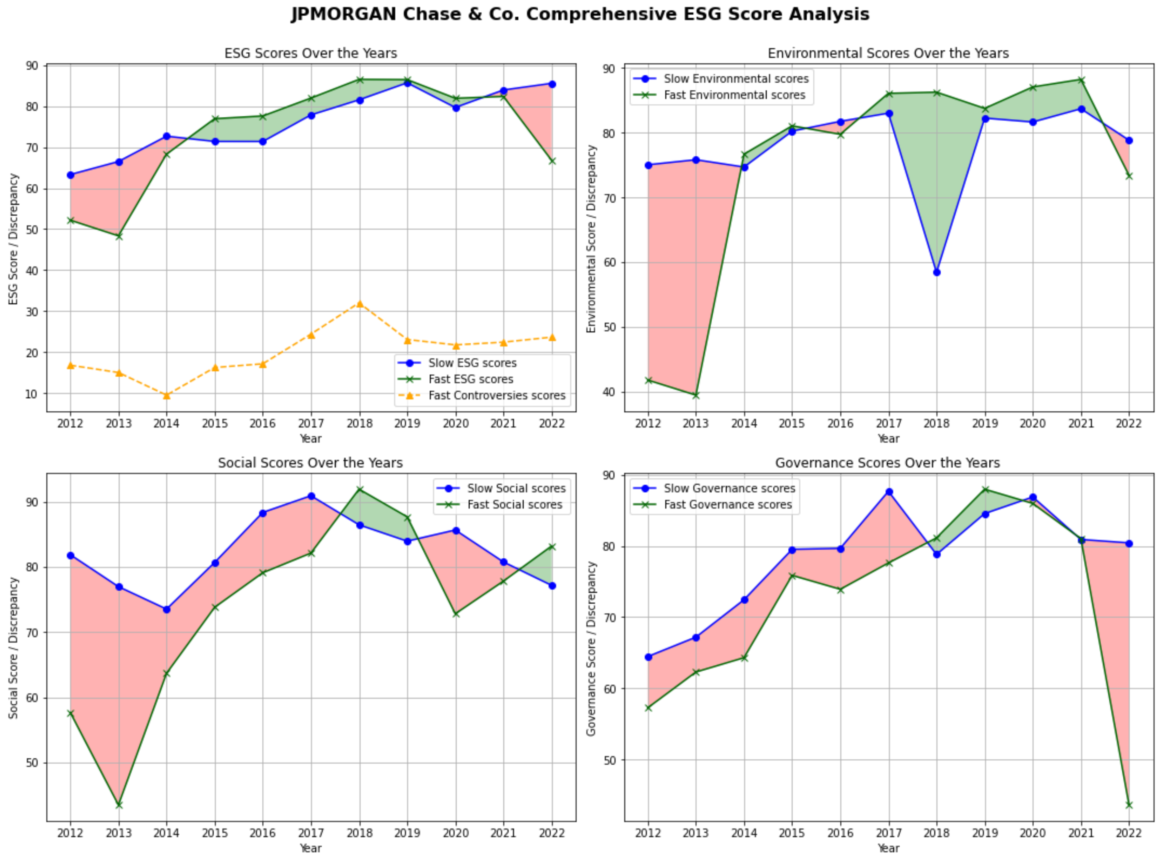

Figure 17: Comprehensive ESG score analysis for JP Morgan Chase & Co.

From 2012 to 2022, JPMorgan Chase & Co.'s ESG analysis largely shows consistent alignment between self-reported scores and media sentiment across Environmental, Social, and Governance components, though significant early discrepancies in ESG, Environmental, and Social scores suggest initial public scepticism. These gaps, where Fast scores were lower, have narrowed over time, indicating better alignment and a reduced risk of greenwashing. However, in 2022, a notable divergence reemerged, particularly in ESG and Governance scores, where Fast scores dropped sharply compared to stable Slow scores, signalling potential greenwashing risks. This decline in Governance scores is largely due to increased regulatory actions, notably the SEC's charge against JPMorgan for inadequate recordkeeping and compliance failures related to the use of personal devices for business communications not preserved as required by law, resulting in a $125 million fine (JPMorgan Chase & Co., 2022). These issues, along with criticisms of oversight and risk management, negatively impacted public perception and underscore the need for robust compliance and ongoing vigilance to align public perceptions with reported data and avoid greenwashing risks.

5.3.2.4 Meta

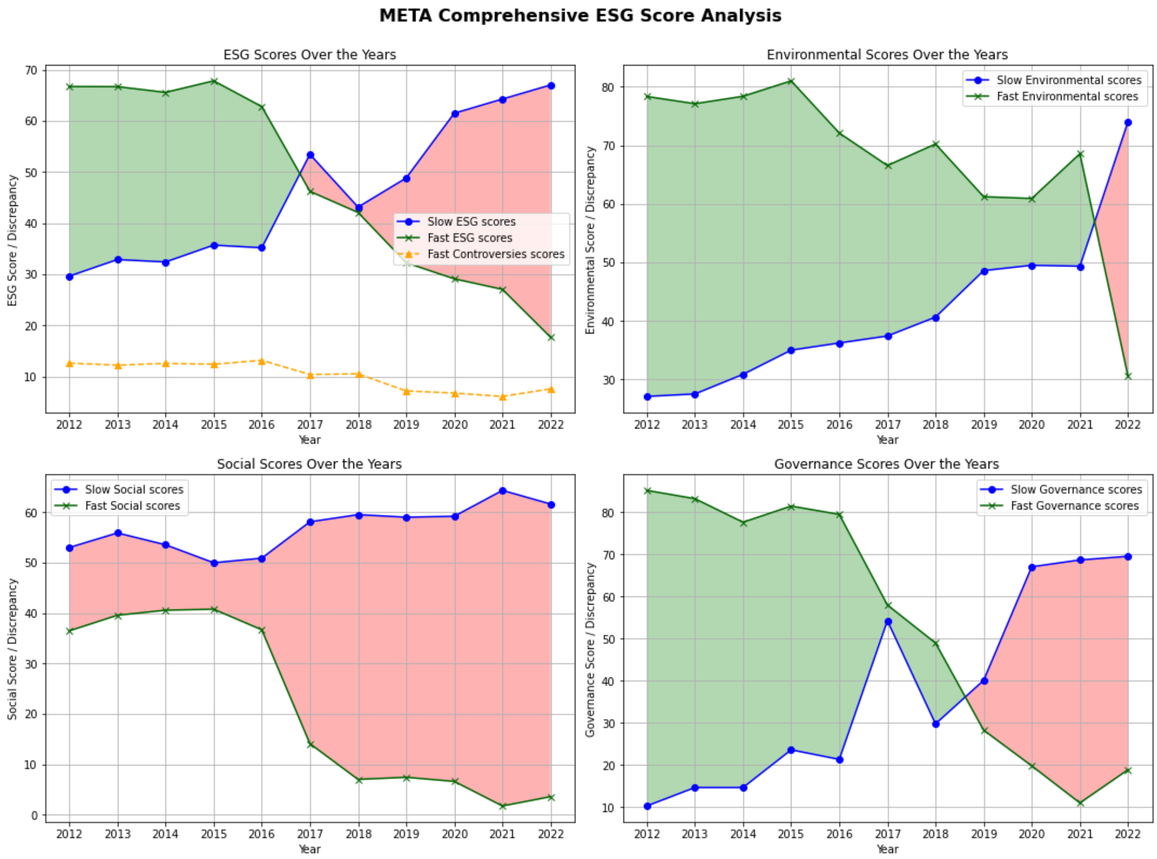

Figure 18: Comprehensive ESG score analysis for META

From 2012 to 2022, Meta's ESG analysis shows a significant misalignment between slow and fast scores, particularly in the Environmental (E) pillar after 2019, with a sharp drop in sentiment in 2022. Despite Meta's achievements in environmental sustainability, such as achieving net zero greenhouse gas emissions supported by 100% renewable energy and setting ambitious goals like becoming water positive by 2030 (Meta, 2023), public perception has suffered due to the broader impacts of their operations and supply chain. Concerns over the environmental impact of the metaverse and Meta's data centres have grown, particularly with the expansion potentially increasing greenhouse gas emissions (Meta, 2023). Additionally, criticism regarding the environmental practices of Meta's suppliers suggests that efforts to decarbonize and enhance supply chain transparency have been inadequate, further impacting public sentiment and widening the gap between reported achievements and perceived sustainability. These factors underscore the challenges Meta faces in reconciling its sustainability ambitions with public expectations and operational realities, leading to discrepancies in ESG scores.

6. Conclusions and recommendations

6.1 Reflection on project aims and objectives

Looking back at the initial aims and objectives detailed in Chapter 2, I believe I have largely accomplished them. I successfully compared and identified correlations between the two data sources for S&P 500 companies. Additionally, I developed a model that can predict potential changes in slow ESG scores based on fast ones. Although the objective of creating a trading strategy using fast scores was adjusted due to methodological and data source limitations, I adapted my approach to provide a basic implementation of a greenwashing indicator. This is equally important in the finance world, offering valuable insights into corporate sustainability practices.

6.2 Reflection on findings and evaluation of process and methodology

The combination of qualitative and quantitative analysis provided a comprehensive understanding of ESG scores and their impact on the financial world. The exploratory data analysis (EDA) and modelling revealed significant insights into the behavior of fast and slow ESG scores for companies. The EDA indicated an overall positive correlation between fast and slow ESG scores over time, with the slow scores aligning more closely with the fast ones through time. Additionally, the EDA showed that within the S&P 500 index, different sectors exhibit varying types of correlations based the ESG scores. Some companies in the sectors steadily improve their ESG scores over time, while others do not. Furthermore, companies with negatively correlated ESG scores tend to receive more media attention compared to those with positive correlations.

The ARIMA predictive model was crucial in addressing the gaps in Refinitiv's data. By providing investors with the ability to monitor potential changes in ESG scores after publication, the model helps to protect investors and their financial positions. Although the hyperparameter optimization of the ARIMA model did not yield optimal results, the parameters still produced decent outcomes and a reasonable results. It is important to note that this model can and should be further optimized with additional data to enhance its accuracy and reliability for investors.

Lastly, the greenwashing indicator I developed, demonstrated its ability to explain and translate certain financial events over time. However, there is still significant room for improvement before it can be used at a professional level. Overall, these results highlight the importance of high-frequency ESG scores, which are nowadays as important as traditional ones.

6.3 Participative qualities of research

The collaborative aspect of the research was crucial to the project's success and methodology. Throughout this project, I engaged with numerous individuals who provided valuable advice that helped me finalize my work. Initially, I interacted with representatives from the helpdesks of Refinitiv and MarketPsych. Their assistance was invaluable in obtaining the data as soon as possible and clarifying specific details that were initially unclear.

At my workplace, where I have been employed for the past two years, two experts made significant contributions to my project. One expert consultant assisted with various technical aspects, such as manipulate the data and helping optimizing my code for data analysis. The other, an expert partner of the company, provided insights into the conceptual aspects of the data. We had many discussions about ESG topics and their importance to companies nowadays.

My university supervisor also played a crucial role, particularly in addressing technical questions related to the modelling part of the project. This collaboration with all these individuals, and our continuous discussions, gave me a deeper understanding of the challenges in the ESG landscape.

These conversations, along with the inevitable frustrations, taught me that roadblocks are part of the process. Without the input from these experts, I might have focused on the wrong objectives, and my research would not have been as complete.

6.4 Recommendations

Based on the project's findings and research, I have the following recommendations:

- **Sentiment data exploration**: While I conducted extensive research on fast ESG scores, I recommend including more sentiment data sources and providers in future studies. This would offer a more comprehensive view of how society's opinions align with companies' self-reported data.
- **Company Regulations**: Understanding how fast and slow ESG scores construct and what they represent, I suggest implementing stricter regulations for companies whose fast ESG scores do not align with their self-reported scores. This is particularly important for companies prone to greenwashing, to ensure transparency and accountability.
- **Controversies Exploration**: Through this research, I gained insights into the significance of ESG scores in the financial landscape. For those interested in studying ESG scores further, I recommend investigating company controversies, such as emissions, CO2 levels, employee satisfaction, and instances of corruption and fraud. This approach will help investors better understand where they are investing their money and be more aware of companies that do not uphold the environmental, social, and governance pillars. Ultimately, investing should not solely focus on financial returns, but also on rewarding companies that respect the environment, treat their employees well, and adhere to sound governance practices.

6.5 Evaluation of personal, organisational and academic development

This research journey significantly enhanced my understanding of ESG scores and their broader impact. Through this thesis, I gained proficiency in handling complex datasets and performing necessary data manipulations to prepare for modelling and analysis. My skills as a data scientist and researcher have been notably enriched.

Additionally, I learned the importance of listening to experts in the field and managing my emotions throughout the research process. Improved communication and collaboration skills enabled me to seek help when needed and persevere through challenges. These experiences have been invaluable in my personal, organizational, and academic development.

7. References, Footnotes and Bibliography

1. Aguilera, R. V., Filatotchev, I., Gospel, H., & Jackson, G. (2008) 'An organizational approach to comparative corporate governance: Costs, contingencies, and complementarities', Organization Science, 19(3), pp. 475–492.

2. Cheng, B., Ioannou, I., & Serafeim, G. (2014) 'Corporate social responsibility and access to finance', Strategic Management Journal, 35(1), pp. 1–23.

3. García-Sánchez, I. M., Rodríguez-Domínguez, L., & Gallego-Álvarez, I. (2020) 'Corporate environmental responsibility and financial performance: A linear and non-linear approach', Business Strategy and the Environment, 29(6), pp. 2701–2715.

4. Pástor, L., Sinha, N., & Swaminathan, B. (2020) 'ESG investing', Journal of Financial Economics, 137(2), pp. 394–413.

5. Dimson, E., Marsh, P., and Staunton, M., 2020. Divergent ESG ratings. The Journal of Portfolio Management. 471. 75–87.

6. Kane, M. J., Price, N., Scotch, M., and Rabinowitz, P., 2014. Comparison of ARIMA and Random Forest time series models for prediction of avian influenza H5N1 outbreaks. BMC bioinformatics. 151. 1–9.

7. Whelan, T., Atz, U., Van Holt, T., and Clark, C., 2021. ESG and financial performance. Uncovering the Relationship by Aggregating Evidence from. 1, 2015–2020.

8. Yu, E. P.-y., Van Luu, B., and Chen, C. H., 2020. Greenwashing in environmental, social and governance disclosures. Research in International Business and Finance. 52, 101192.

9. London Stock Exchange Group (LSEG), 2022. ESG Scores Methodology. [pdf] Available at: https://www.lseg.com.

10. Refinitiv, 2023. ESG Scores Methodology. [pdf] Available at: https://www.refinitiv.com.

11. MarketPsych, 2023. ESG Analytics User Guide. [pdf] Available at: https://www.marketpsych.com.

12. StarMine, 2021. MarketPsych Media Sentiment Model (MMS) – User Guide Version 1.2. [pdf] Available at: https://www.starmine.com.

13. Thomson Reuters, 2022. News Analytics V406 User Guide. [pdf] Available at: https://www.thomsonreuters.com.

14. Durbin, J. and Koopman, S.J., 2012. Time Series Analysis by State Space Methods. Oxford: Oxford University Press. [pdf] Available at: https://global.oup.com/academic.

15. Borovkova, S. and Ying W., 2023. ESG Own Goals Whitepaper.

16. Dewy J., 2023. Sentiment Analysis for Forecasting: Exploring Volatility-at-Risk, Expected Shortfall, and Innovative Trading Strategies.

17. Financial Stability Board, 2023. 2023 TCFD Status Report: Task Force on Climate-related Financial Disclosures. Available at: https://www.fsb.org/2023/10/2023-tcfd-status-report-task-force-on-climate-related-financial-disclosures

18. IR Magazine, 2022. Inside the world of ESG ratings: An academic paper looks at multiple issues in the process and outcomes. Available at: https://www.irmagazine.com/reporting/inside-world-esg-ratings-academic-paper-looks-multiple-issues-process-and-outcomes

19. U.S. Securities and Exchange Commission, 2022. SEC Charges NVIDIA Corporation with Inadequate Disclosures about Impact of Cryptomining. Available at: https://www.sec.gov/news/press-release/2022-79

20. NVIDIA, 2022. Corporate responsibility report. Available at: https://blogs.nvidia.com/blog/corporate-responsibility-report-2

21. JPMorgan Chase & Co., 2022. Annual Report 2022. [pdf] Available at: https://www.jpmorganchase.com/content/dam/jpmc/jpmorgan-chase-and-co/investor-relations/documents/annualreport-2022.pdf

22. Meta, 2023. 2023 Sustainability Report. [pdf] Available at: http://efaidnbmnnnibpcajpcglclefindmkaj/https://sustainability.fb.com/wp-content/uploads/2023/07/Meta-2023-Sustainability-Report-1.pdf

23. Meta, 2022. Meta's progress on combating climate change. Available at: https://about.fb.com/news/2022/11/metas-progress-on-combating-climate-change

24. Covalence SA, 2023. ESG News Collection and Classification Methodology. [pdf] Available at: https://www.covalence.ch

25. Covalence SA, 2020. ESG Ratings Methodology. [pdf] Available at: https://www.covalence.ch

26. OpenAI (2023). ChatGPT (Mar 14 version) [Large language model] Available at: https://chat.openai.com/chat

8. Appendices

8.1 S&P 500 companies

Aflac Inc, AES Corp, Abbott Laboratories, Adobe Inc, Advanced Micro Devices Inc, Air Products and Chemicals Inc, Albemarle Corp, Honeywell International Inc, Allstate Corp, Howmet Aerospace Inc, Hess Corp, Ameren Corp, American Electric Power Company Inc, American Express Co, American International Group Inc, Cencora Inc, AMETEK Inc, Amgen Inc, Amphenol Corp, Analog Devices Inc, Aon PLC, APA Corp (US), Apple Inc, Applied Materials Inc, Archer-Daniels-Midland Co, Atmos Energy Corp, Autodesk Inc, Automatic Data Processing Inc, Autozone Inc, Avalonbay Communities Inc, Avery Dennison Corp, Truist Financial Corp, Baker Hughes Co, Ball Corp, Baxter International Inc, Becton Dickinson and Co, Verizon Communications Inc, W R Berkley Corp, Berkshire Hathaway Inc, Best Buy Co Inc, Bio Rad Laboratories Inc, Boeing Co, Borgwarner Inc, Boston Scientific Corp, Bristol-Myers Squibb Co, Brown-Forman Corp, Cigna Group, CMS Energy Corp, CSX Corp, CVS Health Corp, Coterra Energy Inc, Cadence Design Systems Inc, Camden Property Trust, Campbell Soup Co, Constellation Brands Inc, Capital One Financial Corp, Cardinal Health Inc, Carnival Corp, Caterpillar Inc, JPMorgan Chase & Co, Chevron Corp, Church & Dwight Co Inc, Cincinnati Financial Corp, Cisco Systems Inc, Cintas Corp, Clorox Co, Coca-Cola Co, Colgate-Palmolive Co, Comerica Inc, Conagra Brands Inc, Consolidated Edison Inc, Cooper Companies Inc, Molson Coors Beverage Co, Copart Inc, Corning Inc, Cummins Inc, DR Horton Inc, DTE Energy Co, Danaher Corp, Darden Restaurants Inc, Target Corp, Deckers Outdoor Corp, Deere & Co, Walt Disney Co, Dollar Tree Inc, Dominion Energy Inc, Dover Corp, Duke Energy Corp, Eastman Chemical Co, Eaton Corporation PLC, Ecolab Inc, Edison International, Electronic Arts Inc, Emerson Electric Co, EOG Resources Inc, Entergy Corp, Equifax Inc, EQT Corp, Equity Residential, Everest Group Ltd, Expeditors International of Washington Inc, Exxon Mobil Corp, FMC Corp, Nextera Energy Inc, Fair Isaac Corp, Fastenal Co, FedEx Corp, Fifth Third Bancorp, Fiserv Inc, FirstEnergy Corp, Franklin Resources Inc, Freeport-McMoRan Inc, Arthur J Gallagher & Co, Gartner Inc, General Dynamics Corp, General Electric Co, General Mills Inc, Genuine Parts Co, Gilead Sciences Inc, WW Grainger Inc, Halliburton Co, L3Harris Technologies Inc, Hartford Financial Services Group Inc, Hasbro Inc, Healthpeak Properties Inc, Jack Henry & Associates Inc, Hershey Co, HP Inc, Hologic Inc, Home Depot Inc, Hormel Foods Corp, Host Hotels & Resorts Inc, Hubbell Inc, Humana Inc, J B Hunt Transport Services Inc, Huntington Bancshares Inc, Biogen Inc, Mosaic Co, IDEX Corp, IDEXX Laboratories Inc, Illinois Tool Works Inc, Incyte Corp, Trane Technologies PLC, Intel Corp, International Business Machines Corp, International Flavors & Fragrances Inc, International Paper Co, Interpublic Group of Companies Inc, Intuit Inc, Jabil Inc, Jacobs Solutions Inc, Johnson & Johnson, KLA Corp, Kellanova, KeyCorp, Kimberly-Clark Corp, Kroger Co, Laboratory Corporation of America Holdings, Lam Research Corp, Estee Lauder Companies Inc, Lennar Corp, Eli Lilly and Co, Bath & Body Works Inc, Lockheed Martin Corp, Loews Corp, Lowe's Companies Inc, M&T Bank Corp, MGM Resorts International, Marsh & McLennan Companies Inc, Marriott International Inc, Martin Marietta Materials Inc, Masco Corp, McCormick & Company Inc, McDonald's Corp, S&P Global Inc, Mckesson Corp, Medtronic PLC, Bank of New York Mellon Corp, Microsoft Corp, Microchip Technology Inc, Micron Technology Inc, Mid-America Apartment Communities Inc, 3M Co, Mohawk Industries Inc, Morgan Stanley, Motorola Solutions Inc, Viatris Inc, NVR Inc, NetApp Inc, Newmont Corporation, Nike Inc, Nordson Corp, Norfolk Southern Corp, Eversource Energy, Xcel Energy Inc, Northern Trust Corp, Northrop Grumman Corp, Wells Fargo & Co, Nucor Corp, Occidental Petroleum Corp, Old Dominion Freight Line Inc, Omnicom Group Inc, ONEOK Inc, Oracle Corp, O'Reilly Automotive Inc, Exelon Corp, PG&E Corp, PNC Financial Services Group Inc, PPL Corp, PPG Industries Inc, Paccar Inc, PTC Inc, Parker-Hannifin Corp, Paychex Inc, Pentair PLC, PepsiCo Inc, Pfizer Inc, Altria Group Inc, Conocophillips, Pinnacle West Capital Corp, Pioneer Natural Resources Co, T Rowe Price Group Inc, Procter & Gamble Co, Progressive Corp, Public Service Enterprise Group Inc, Pultegroup Inc, Qualcomm Inc, Quanta Services Inc, Raymond James Financial Inc, Realty Income Corp, Regeneron Pharmaceuticals Inc, Regency Centers Corp, Resmed Inc, Arch Capital Group Ltd, Robert Half Inc, Rockwell Automation Inc, Rollins Inc, Roper Technologies Inc, Ross Stores Inc, AT&T Inc, Travelers Companies Inc, Henry Schein Inc, Schlumberger NV, Charles Schwab Corp, Sempra, Sherwin-Williams Co, Simon Property Group Inc, A O Smith Corp, Snap-On Inc, Southern Co, Southwest Airlines Co, Stanley Black & Decker Inc, US Bancorp, Starbucks Corp, State Street Corp, Stryker Corp, Gen Digital Inc, Synopsys Inc, Sysco Corp, TJX Companies Inc, Bio-Techne Corp, Teleflex Inc, Teradyne Inc, Texas Instruments Inc, Textron Inc, Thermo Fisher Scientific Inc, Globe Life Inc, DaVita Inc, Tractor Supply Co, Citigroup Inc, Yum! Brands Inc, Trimble Inc, Tyson Foods Inc, Marathon Oil Corp, Waste Management Inc, Union Pacific Corp, UDR Inc, UnitedHealth Group Inc, RTX Corp, Universal Health Services Inc, Valero Energy Corp, Ventas Inc, VeriSign, Inc, Vertex Pharmaceuticals Inc, Vulcan Materials Co, Walmart Inc, Walgreens Boots Alliance Inc, Waters Corp, Western Digital Corp, Evergy Inc, Westinghouse Air Brake Technologies Corp, Weyerhaeuser Co, Williams Companies Inc, WEC Energy Group Inc, Zebra Technologies Corp, Chubb Ltd, Johnson Controls International PLC, Royal Caribbean Cruises Ltd, American Tower Corp, Moody's Corp, Quest Diagnostics Inc, Steel Dynamics Inc, Factset Research Systems Inc, Prologis Inc, United Rentals Inc, CenterPoint Energy Inc, Boston Properties Inc, Essex Property Trust Inc, Alexandria Real Estate Equities Inc, CH Robinson Worldwide Inc, Mettler-Toledo International Inc, Iron Mountain Inc, West Pharmaceutical Services Inc, NiSource Inc, Skyworks Solutions Inc, Bank of America Corp, Amazon.com Inc, Brown & Brown Inc, Ralph Lauren Corp, Alliant Energy Corp, Tyler Technologies Inc, Cognizant Technology Solutions Corp, Crown Castle Inc, eBay Inc, Goldman Sachs Group Inc, NVIDIA Corp, Booking Holdings Inc, Republic Services Inc, Pool Corp, CoStar Group Inc, Costco Wholesale Corp, Devon Energy Corp, Revvity Inc, Take-Two Interactive Software Inc, Akamai Technologies Inc, Teledyne Technologies Inc, United Parcel Service Inc, Juniper Networks Inc, Edwards Lifesciences Corp, Agilent Technologies Inc, BlackRock Inc, F5 Inc, MetLife Inc, Packaging Corp of America, SBA Communications Corp, ON Semiconductor Corp, Ford Motor Co, Global Payments Inc, Tapestry Inc, Align Technology Inc, ANSYS Inc, Charles River Laboratories International Inc, Carmax Inc, Illumina Inc, Intuitive Surgical Inc, Fidelity National Information Services Inc, Zimmer Biomet Holdings Inc, Accenture PLC, Elevance Health Inc, Invesco Ltd, Equinix Inc, Garmin Ltd, Monster Beverage Corp, Mondelez International Inc, Principal Financial Group Inc, Axon Enterprise Inc, Willis Towers Watson PLC, Centene Corp, Prudential Financial Inc, Netflix Inc, J M Smucker Co, Nasdaq Inc, Wynn Resorts Ltd, Comcast Corp, CME Group Inc, Seagate Technology Holdings PLC, Molina Healthcare Inc, LKQ Corp, NRG Energy Inc, Assurant Inc, CBRE Group Inc, Salesforce Inc, Regions Financial Corp, Domino's Pizza Inc, T-Mobile US Inc, Extra Space Storage Inc, Digital Realty Trust Inc, Marketaxess Holdings Inc, Monolithic Power Systems Inc, Las Vegas Sands Corp, Celanese Corp, Dexcom Inc, Builders FirstSource Inc, Expedia Group Inc, CF Industries Holdings Inc, Ameriprise Financial Inc, Mastercard Inc, Intercontinental Exchange Inc, Leidos Holdings Inc, Paramount Global, Live Nation Entertainment Inc, Chipotle Mexican Grill Inc, United Airlines Holdings Inc, TransDigm Group Inc, First Solar Inc, Broadridge Financial Solutions Inc, Delta Air Lines Inc, Insulet Corp, TE Connectivity Ltd, Discover Financial Services, Blackstone Inc, Lululemon Athletica Inc, Ulta Beauty Inc, MSCI Inc, Philip Morris International Inc, Visa Inc, American Water Works Company Inc, Keurig Dr Pepper Inc, Warner Bros Discovery Inc, Broadcom Inc, Verisk Analytics Inc, Merck & Co Inc, Dollar General Corp, Fortinet Inc, Charter Communications Inc, Generac Holdings Inc, LyondellBasell Industries NV, Cboe Global Markets Inc, Tesla Inc, NXP Semiconductors NV, Targa Resources Corp, Kinder Morgan Inc, HCA Healthcare Inc, General Motors Co, Huntington Ingalls Industries Inc, Marathon Petroleum Corp, Xylem Inc, Aptiv PLC, Epam Systems Inc, Enphase Energy Inc, Corpay Inc, Norwegian Cruise Line Holdings Ltd, Meta Platforms Inc, Diamondback Energy Inc, ServiceNow Inc, Phillips 66, Palo Alto Networks Inc, AbbVie Inc, Zoetis Inc, IQVIA Holdings Inc, News Corp, CDW Corp, Allegion PLC, Hilton

Worldwide Holdings Inc, American Airlines Group Inc, Alphabet Inc, Arista Networks Inc, Paycom Software Inc, Catalent Inc, Synchrony Financial, Citizens Financial Group Inc, Caesars Entertainment Inc, Keysight Technologies Inc, Qorvo Inc, ETSY Inc, WestRock Co, Kraft Heinz Co, PayPal Holdings Inc, Hewlett Packard Enterprise Co, Match Group Inc, STERIS plc, Fortive Corp, Lamb Weston Holdings Inc, Invitation Homes Inc, Ingersoll Rand Inc, Dupont De Nemours Inc, VICI Properties Inc, Dayforce Inc, Moderna Inc, Fox Corp, Dow Inc, Uber Technologies Inc, Corteva Inc, Amcor PLC, Super Micro Computer Inc, Carrier Global Corp, Otis Worldwide Corp, Airbnb Inc, Constellation Energy Corp, GE Healthcare Technologies Inc, Kenvue Inc, Veralto Corp, GE Vernova Inc.

8.2 Categories table

Category	Explanation
CSR Strategy	Reflects a company's sustainability reporting
Community	Measures the company's commitment towards being a good citizen, protecting public health and respecting business ethics
Emissions	Measures a company's commitment and effectiveness towards reducing environmental emissions in the production and operational processes
Environmental Innovation	Reflects a company's reduction of environmental impact and the creation of new market opportunities through new green technologies and design
Human Rights	Measures a company's effectiveness at respecting fundamental human rights
Management	Measures a company's commitment and effectiveness towards following best practice corporate governance principles
Product Responsibility	Reflects a company's capacity to produce quality goods and services honoring the customer's health and safety, integrity and data privacy
Resource Use	Reflects a company's energy efficiency and supply chain sustainability

Shareholders	Measures a company's effectiveness towards equal treatment of shareholders and the use of anti-takeover devices
Workforce	Measures job satisfaction, workplace safety, maintaining diversity and equal opportunities, and development opportunities for the company's workforce

Table 9: Fast ESG categories

8.3 Correlation categories between Fast and Slow ESG scores

Figure 19: High correlation between ESG scores

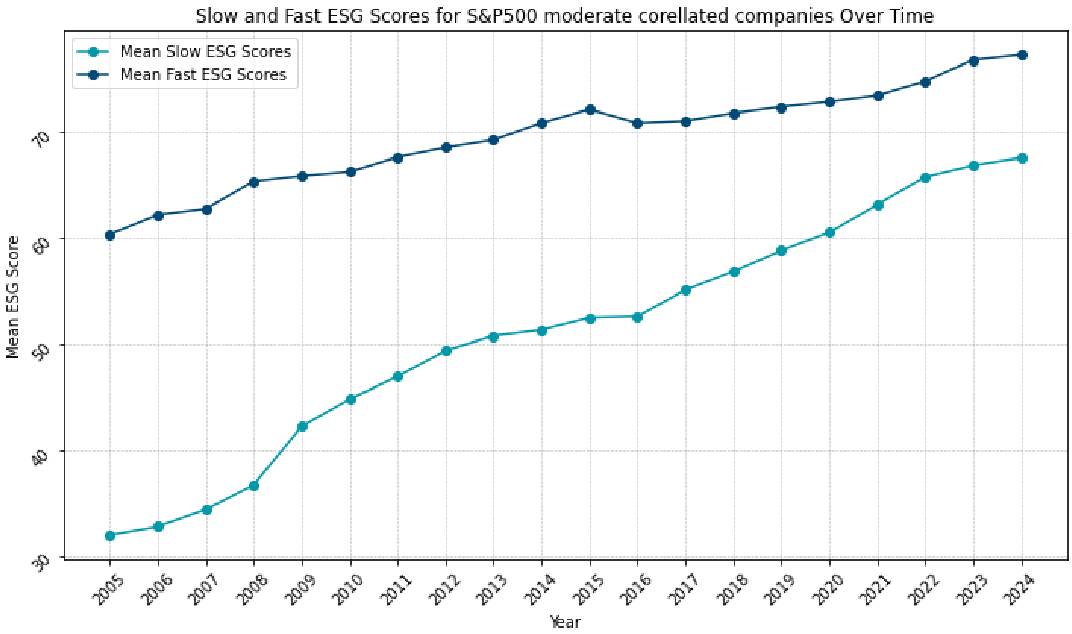

Figure 20: Moderate correlation between ESG scores

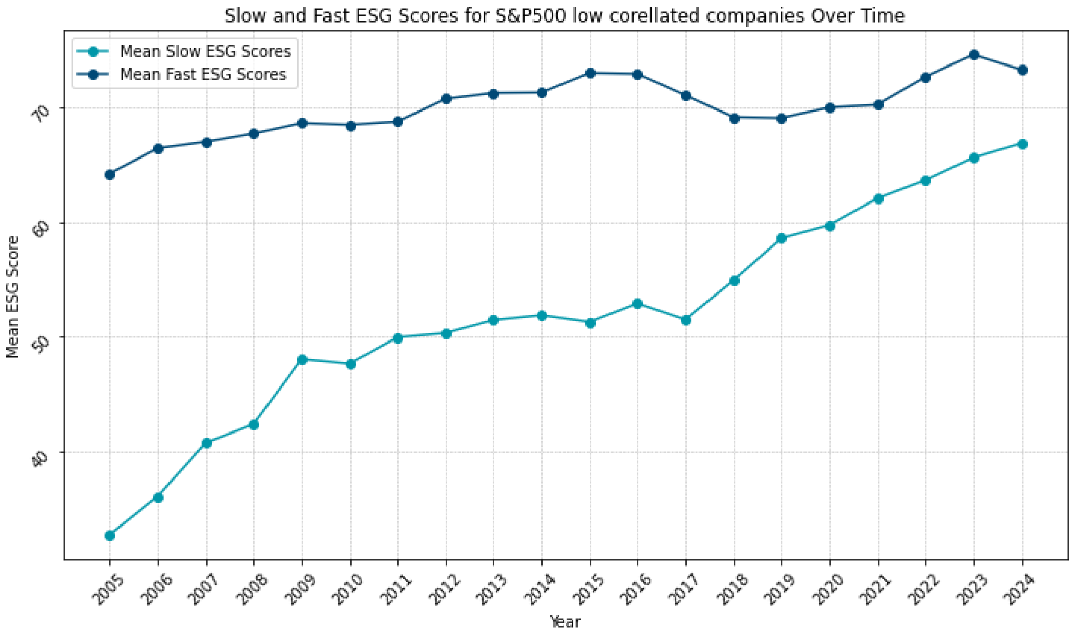

Figure 21: Low correlation between ESG scores

8.4 Overview of Industry Categories by Correlation Group

Industry Category / Type of correlation	High	Moderate	Low	Negative
Communications & Networking	3	2	1	2
Insurance	3	13	4	1
Financial Technology (Fintech) & Infrastructure	2	0	1	0
Freight & Logistics Services	1	4	2	1
Investment Banking & Investment Services	2	10	4	2
Natural Gas Utilities	1	1	0	0
Residential & Commercial REITs	8	10	2	2
Food & Drug Retailing	3	1	2	0
Household Goods	1	0	0	0
Personal & Household Products & Services	4	3	0	0
Real Estate Operations	2	0	0	1
Consumer Goods Conglomerates	0	2	2	1
Construction Materials	0	1	1	0
Leisure Products	0	1	2	0
Healthcare Equipment & Supplies	12	9	4	2
Biotechnology & Medical Research	4	0	1	3
Healthcare Providers & Services	1	7	2	0
Multiline Utilities	2	3	1	0
Diversified Retail	1	4	1	0
Telecommunications Services	0	0	2	2
Machinery, Tools, Heavy Vehicles, Trains & Ships	7	14	5	6
Professional & Commercial Services	11	9	3	1
Hotels & Entertainment Services	3	7	3	3
Computers, Phones & Household Electronics	0	2	2	3
Office Equipment	0	1	1	0
Media & Publishing	1	2	2	3
Beverages	1	4	1	1

Water & Related Utilities	0	1	0	0
Oil & Gas	5	4	4	2
Construction & Engineering	1	1	0	0
Textiles & Apparel	3	0	0	1
Homebuilding & Construction Supplies	0	3	1	0
Food & Tobacco	4	8	0	4
Automobiles & Auto Parts	3	2	0	1
Pharmaceuticals	4	2	3	4
Metals & Mining	1	3	1	1
Software & IT Services	15	17	6	5
Renewable Energy	1	1	0	0
Specialty Retailers	3	5	2	1
Chemicals	3	5	3	3
Electric Utilities & IPPs	9	7	1	2
Banking Services	2	8	3	5
Oil & Gas Related Equipment and Services	1	4	2	0
Passenger Transportation Services	1	0	0	2
Containers & Packaging	1	4	0	1
Aerospace & Defense	1	7	1	2
Semiconductors & Semiconductor Equipment	3	12	2	3
Electronic Equipment & Parts	3	3	1	0
Unknown	6	6	3	2

Table 10: Analytically Industry Categories Overview by Correlation Group

8.5 Sectors and subsectors

Sector	Subsector
Communication Services	Communications & Networking, Telecommunications Services, Media & Publishing
Consumer Discretionary	Leisure Products, Hotels & Entertainment Services, Diversified Retail, Textiles & Apparel, Automobiles & Auto Parts, Homebuilding & Construction Supplies, Specialty Retailers, Passenger Transportation Services, Consumer Goods Conglomerates
Consumer Staples	Food & Drug Retailing, Beverages, Food & Tobacco, Household Goods, Personal & Household Products & Services
Energy	Oil & Gas, Oil & Gas Related Equipment and Services, Renewable Energy
Financials	Insurance, Financial Technology (Fintech) & Infrastructure, Investment Banking & Investment Services, Banking Services
Healthcare	Healthcare Equipment & Supplies, Biotechnology & Medical Research, Healthcare Providers & Services, Pharmaceuticals
Industrials	Freight & Logistics Services, Machinery, Tools, Heavy Vehicles, Trains & Ships, Professional & Commercial Services, Construction & Engineering, Containers & Packaging, Aerospace & Defense
Information Technology	Software & IT Services, Computers, Phones & Household Electronics, Office Equipment, Semiconductors & Semiconductor Equipment, Electronic Equipment & Parts
Materials	Construction Materials, Chemicals, Metals & Mining
Real Estate	Residential & Commercial REITs, Real Estate Operations
Utilities	Natural Gas Utilities, Electric Utilities & IPPs, Multiline Utilities, Water & Related Utilities

Table 11: Sectors and subsectors of S&P500