Louvain School of Management

The Influence of Media-Driven Investor Attention on Clean Technology Investment

Investigating the Role of News Sentiment and Political Orientation

Author: Clara Verdonck
Supervisor: Prabal Shrestha

Academic year 2024-2025

Dissertation for the master of

CEMS Master in International Management

AI Declaration

During the preparation of this master’s thesis, the author utilized ChatGPT for the following purpose:

- Paraphrasing original text written by the author in the literature review to improve clarity and academic style.
- Generating and debugging R code and Python code to support data analysis and econometric modeling.

After using ChatGPT, the author reviewed and edited the content produced by the tool. Full responsibility is taken for the final content presented in this thesis.

By signing this declaration, the author affirms that the content of this master’s thesis reflects original work, augmented by the responsible use of AI.

Clara Verdonck

26/05/2025

Acknowledgements

I thank Manuela Verdonck and Nancy Verdonck for proofreading this thesis and for their support throughout the process.

Supervisor: Prabal Shrestha

Abstract :

This thesis examines how investor attention – proxied by legal newspaper coverage of clean technology news articles – influences stock returns of clean technology companies. Focusing on 108 U.S. and UK listed firms across hydrogen, ammonia, methanol, e-fuels, and electric vehicle sectors, it tests five hypotheses linking overall news coverage effect, sentiment (positive/neutral/negative), political alignment (left/center/right/far-right), their interaction, and technology-match effects to cumulative abnormal returns (CARs). Employing event-study methodology over 2007–2022 to compute 10-day CARs via a CAPM framework, followed by fixed effects panel regressions, the analysis reveals that both positive and negative tones enhance short-term CARs relative to neutral coverage, while news outlets at the political extremes (especially far-right and left) amplify these effects. Technology-match analysis demonstrates that methanol and hydrogen companies benefit significantly when news specifically addresses their core technology, and methanol coverage generates spillover effects across sectors. These findings underscore media coverage and content as small, but significant factors in influencing clean technology stock returns. Future research could extend this framework by examining investor attention effects in markets outside the U.S. and UK, and by integrating other attention measures (e.g. social media or search data) to capture real-time dynamics.

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
Louvain School of Management

Place des Doyens, 1 bte L2.01.01, 1348 Louvain-la-Neuve
Boulevard Emile Devreux 6, 6000 Charleroi, Belgique
Chaussée de Binche 151, 7000 Mons, Belgique

www.uclouvain.be/lsm

Table of Contents

1. Introduction

Ambitious net-zero targets define the global political landscape, with the EU Green Deal aiming for climate neutrality by 2050, the U.S. targeting 100% carbon-free electricity by 2035, and China's Five-Year Plan committing to net-zero by 2060 (European Commission, 2023; U.S. Department of State, n.d.; Hepburn et al., 2021). Clean technologies have become indispensable for decarbonization, yet their markets are inherently complex: they are capital-intensive, policy-dependent, and often subject to long payback periods (Jain, 2024). In such uncertain contexts, the behavioral aspect of investment patterns, what drives adoption and how, is interesting to examine. In particular, investor attention (i.e. the limited attention of investors when having to choose between particular events, companies, stocks or industries (Maula & Lukkarinen, 2022)) emerges as a critical, yet underexplored, determinant of clean technology investment and, by extension, adoption (Barber & Odean, 2008; Merton, 1987; Kahneman, 1973). The way clean technologies are portrayed, positively or negatively, through left- or right-leaning news outlets, can affect which companies attract investment, and how quickly they scale.

This thesis investigates a central question: '*How does investor attention to clean technology news coverage, conditioned by media tone and political orientation, influence the investment into clean technology alternatives?*' Linking insights from innovation diffusion theory (Rogers, 2003), clean tech adoption frameworks (Kemp & Volpi, 2008), and behavioral finance (Maula & Lukkarinen, 2022), this study analyzes how media coverage impacts the stock market performance of clean technology companies.

Regarding this question, five hypotheses are tested, concerning (1) the overall coverage effect (2) tone effects, (3) political alignment effects, (4) their interaction, and (5) technology-match effects. To address this question and the hypotheses, an empirical study is conducted that combines event study methodology with fixed effects panel regression analysis. Investor attention is proxied by media coverage (Cumming et al., 2016), drawing on a uniquely compiled dataset of over 1500 clean technology newspaper headlines from 2007 to 2022, categorized by sentiment and political alignment, and matched with stock market data from 108 U.S. and UK-based clean tech companies. The study examines the impact of media coverage on stock market performance, both across the entire sample (PART 1) and within specific clean technology sectors (PART 2): hydrogen, ammonia, methanol, e-fuels, and electric vehicles. By analyzing cumulative abnormal returns (CARs) following media events, the research assesses how variations in media tone and political orientation influence investor reactions. Further methodological details, including variable operationalization, event window selection, and robustness checks, are provided in the methodology chapter.

The findings reveal that media coverage significantly influences investor behavior, but not uniformly. In the broad panel regression analysis (PART 1.2), extreme emotions – positive or negative sentiment – and extreme political orientations – left and far right – generally elicit a more positive CAR than the neutral sentiment and centrist politics alternative. In PART 2, the results differ per technology. Methanol and hydrogen companies benefit most from news that is specifically about methanol or hydrogen, respectively. Furthermore, news about methanol benefits all five technologies. Similarly to the results in PART 1.2, the panel regression results in PART 2.2 show that more extreme emotions and politics elicit stronger returns, which can be both positive (e.g.

hydrogen news articles with positive sentiment from right-wing newspapers have a positive effect on the CAR of hydrogen companies) or negative (e.g. EV news articles with positive sentiment from right-wing newspapers have a negative effect on the CAR of EV companies). This shows that results are not uniform and that negative sentiment, for instance, does not necessarily give rise to a negative effect on CARs, as opposed to the findings in previous literature (Tetlock, 2007).

In conducting this research, several challenges arose. The complexity of the research design – which integrates event-study methodology, fixed-effects panel regressions, sentiment analysis, political-bias classification, and technology-match subsetting – demanded careful coordination and rigor. Second, data availability proved to be a constraint: while over 1500 headlines were ultimately compiled, the uneven distribution of articles across technologies and news outlets limited statistical power in some subgroup analyses (notably ammonia, methanol, and e-fuels). Finally, addressing extreme observations in CARs necessitated winsorization, underscoring the sensitivity of financial event studies to outliers.

This thesis makes several contributions. It is among the first that integrates sentiment and political-bias from legal newspapers into an event-study and panel-regression analysis for clean technology. This thesis empirically documents technology-specific heterogeneity, revealing that coverage of emerging sectors (hydrogen, methanol) yields significantly larger CARs than for mature sectors (electric vehicles). This underlines that media coverage can be crucial for adoption of emerging technologies. For investors and companies, the research provides guidance on how strategic media engagement might enhance capital access. For policymakers, they underscore the need to broaden coverage of under-represented clean technology domains to encourage investment; and for researchers, they open interesting ground for exploring how media-driven attention dynamics interact with financial market structures to shape sustainable investment outcomes.

Following this introduction, the thesis proceeds as follows. Chapter 2 reviews the relevant literature on clean technology, investor attention, and media effects. Chapter 3 develops the main hypotheses. Chapter 4 outlines the data and research design. Chapter 5 presents the empirical analysis and results. Finally, Chapter 6 concludes, discusses limitations, and suggests directions for future research.

2. Literature Review

2.1 Clean Technology

2.1.1 Global Climate Goals and a Need for More Clean Technology Adoption

The global regulatory landscape is increasingly focused on achieving net-zero emissions to meet critical climate goals. To this end, there is consensus among organizations, governments and expert bodies that the adoption of clean technology is crucial. The European Climate Neutrality Observatory describes clean technologies as the backbone of a decarbonized economy (ECNO, n.d.). Adoption or diffusion refers to the decision to use and implement a new idea (Kemp and Volpi, 2008), while the "rate of adoption is the relative speed with which an innovation is adopted by members of a social system" (Rogers, 2003).

In the United States (U.S.), the Biden administration has set federal climate goals of cutting U.S. greenhouse gas emissions 61-66% by 2035 relative to 2005 levels, 50% electric vehicle sales share by 2030, 100% carbon pollution-free electricity by 2035 and a net-zero emissions economy by 2050 (U.S. Department of State, n.d.). In the European Union (EU), the European Climate Law writes into law the goal of the European Green Deal of Europe becoming climate-neutral by 2050, with the intermediary goal of reducing greenhouse gas emissions by minimum 55% by 2030, compared to 1990 levels (European Commission, n.d.). Clean technologies are at the heart of the Green Deal Industrial Plan and NextGenerationEU, which aim to enhance Europe's competitiveness in the net-zero industry by simplifying regulation, by providing faster access to funding, by enhancing skills and by opening up trade (European Commission, 2023). In China, its 14th Five-Year-Plan for the period 2021-2025, commits to achieving carbon neutrality by 2060. Its green transition will involve structural transformation of industry towards higher skills with less material input, new technologies and re-designing energy and transport systems (Hepburn et al., 2021).

2.1.2 Overview of Clean Technology and Reasons for Adoption

Clean technology generally refers to new technologies and business models that aim to reduce the environmental footprint and to minimize environmental pollution. Over the past ten years, clean technology has undergone transformations that have reshaped energy production, waste management and infrastructure development. Clean technologies must achieve several attributes to fulfill their purpose: raw material conservation; optimal use of raw materials and other materials; optimal use of other valuable ingredients, namely, energy and water; optimization of production processes; safe disposal of unwanted waste; maximum possible recycling of unavoidable waste; prevention of accidents; prevention of pollution by risk management (Muthu, 2020).

Clean technology plays a big role in facilitating the green transition by enabling systemic decarbonization across multiple sectors (Jain, 2024). Advances in renewable energy, particularly solar and wind, combined with energy storage solutions, have significantly improved the reliability and scalability of low-carbon power systems. In the transportation sector, the electrification of vehicles, alongside the adoption of alternative fuels and smart mobility solutions, contributes to substantial reductions in greenhouse gas emissions. Clean technologies also support the development of sustainable infrastructure through energy-efficient buildings and green construction practices. Beyond environmental benefits, clean technology contributes to broader socioeconomic development by creating green jobs, improving public health, and aligning industrial practices with the Sustainable Development Goals (United Nations, 2015).

The clean energy financial market emerges as the fastest growing energy segment due to a high increase in investment (Ghosh & Jana, 2024). The global market size for clean energy technologies on the supply side reached almost $790 billion in the end of Q1 2023 (Statista, 2025b). The total annual global investment in clean energy transition technologies exceeded $2.1 trillion in 2024, compared to only $231 billion in 2010 (Statista, 2025a). The electrified transport sector ($757 billion), the renewable energy sector ($728 billion) and the power grid sector ($390 billion) had the highest investment value overall among energy transition-relevant sectors worldwide in 2023. In contrast, the other seven sectors combined (hydrogen, carbon capture and storage, nuclear, clean shipping, electrified heat, and clean industry) accounted for only 7.4% of total investment and saw a 23% investment decline, highlighting the persistent difficulties in expanding these emerging clean technologies (Bloomberg, 2025).

The transportation sector represents a big share of the overall greenhouse gas (GHG) emissions. In 2020, road transport from passenger cars, trucks and buses represented about 24% of EU's total GHG (European Commission, 2021). Maritime transport represented 4% of the EU's total GHG emissions in 2021 (European Commission, 2022). Therefore, transport GHG emissions amount to almost 30% of total GHG emissions. The numbers in the U.S. show similar results with 27% of GHG coming from the transportation sector. The transport sector is also the largest consumer of oil and is also the fastest growing in terms of energy consumption (International Energy Agency, 2018).

A particular challenge lies in decarbonizing hard-to-abate sectors, meaning those industries that experience a particular challenge in reducing the reliance on traditional, energy dense fuels that are often high in carbon (World Economic Forum, 2024). The necessity of decarbonizing hard-to-abate sectors, such as shipping, aviation, power generation, and other industries has created a growing interest for e-fuels. E-fuels are synthetic fuels produced from renewable hydrogen and (if carbon-based) captured CO_2. Examples include e-methanol and e-ammonia, which have gained attention as zero- or low-carbon alternatives for hard-to-electrify sectors. Both fuels can be produced using renewable electricity (via hydrogen), offering pathways to reduce greenhouse gas emissions in sectors that historically rely on fossil fuels. However, uptake has been gradual – renewable hydrogen and e-fuels are still used in only small quantities today (International Energy Agency, 2024) – due to uncertainty regarding the path to take forward. For instance, the Platform for Electromobility considers the risks related to e-fuels, such as higher costs, regulatory uncertainty and health risks (Platform for Electromobility, 2025).

Among e-fuels, methanol and ammonia are often compared to each other and can be seen as competitors or complementary solutions, depending on the sector. In the maritime industry, methanol and ammonia are in direct competition, with methanol often seen as an intermediary solution, while ammonia is the ultimate solution on the long term (IMO, n.d.; IMO, 2023; IRENA, 2021; S&P Global, 2025). Outside the shipping industry, competition is found to be less direct because the optimal use cases often diverge. In aviation, ammonia would be preferred as a hydrogen carrier for a potential future aircraft, while methanol would more likely contribute indirectly via conversion. In energy storage and transport, both e-fuels could complement each other. Ammonia is a gas, but more energy dense, whereas methanol is a liquid and could therefore be used in applications where carbon-based liquid is needed (e.g. vehicles) (World Bank, 2024).

2.1.3 Factors Driving Adoption of Clean Technology

Adoption Process

In analyzing the factors driving the adoption of innovations, the study by Rogers (2003) works out a theoretical framework, the innovation-decision process, which is the process through which an individual acquires (1) knowledge about the innovation, (2) a persuasion or opinion about the innovation, (3) a decision of whether to adopt of reject the idea. This then evolves into (4) an implementation of the innovation, and lastly (5) a confirmation of the innovation. During stage (2), Roger identifies five perceived attributes of innovations that drive the rate of adoption and guide the potential adopter into adoption or rejection of an innovation:

1. Relative Advantage is the degree to which an innovation is perceived as being better than an alternative. It is often expressed in economic profitability.
2. Compatibility is defined as the degree to which an innovation is seen as consistent with the existing values, past experiences, and needs of a potential adopter. This idea is also related to familiarity and trust in past innovations.
3. Complexity reflects the degree to which an innovation is perceived as complex to use and understand. Complexity is the only attribute of the five that is negatively related to the rate of adoption of an innovation.
4. Trialability is the degree to which an innovation can be experimented with on a limited basis. A triable innovation is associated with lower uncertainty.
5. Observability refers to the visibility of the results of an innovation to others.

Kemp and Volpi (2008) provide further insight into the diffusion of clean technologies specifically. A central contribution of their study is the recognition that companies face multiple technological pathways when responding to environmental pressures. They distinguish between three categories of response: (1) end-of-pipe solutions, (2) process adaptations, and (3) full process substitutions. This highlights that clean technologies diffuse in a competitive landscape, where the adoption of one technology can happen at the expense of the other. The authors underscore that economic attractiveness is a necessary, but insufficient condition for adoption. Innovations perceived as economically superior can still diffuse slowly due to sunk costs or risk aversion. Adoption is further influenced by the endogenous nature of learning – i.e. as more companies adopt the innovation, knowledge about it increases, thereby reinforcing further adoption. A factor hampering adoption of an innovation is information asymmetry, which can be reduced through communication. Rogers (2003) finds that mass media channels are useful to inform about less complex innovations, while interpersonal contact is more important for innovations that are more complex. If mass media channels are used to communicate about complex ideas, then the rate of adoption is slower.

Policy and Regulation

As mentioned, policy and regulation is a driving factor of adoption (European Commission, 2023; U.S. Department of State, n.d.; Hepburn et al., 2021), while policy uncertainty hampers it. Global economic and policy uncertainty shocks reduce renewable energy patent filings by 1.5%, with stronger effects during recessions (Bettarelli et al., 2024). The Organisation for Economic

Cooperation and Development (OECD) identifies policy uncertainty as a key deterrent, delaying capital-intensive green investments due to irreversibility concerns (OECD, 2022).

Stock Market Investment

Stock market investment in clean technology is another important adoption factor. Empirical research indicates that well-functioning financial markets can accelerate the diffusion and adoption of new technologies. Rising stock prices, greater market liquidity, or broader market development often provide companies with easier access to capital needed for innovation and adoption. A cross-country study by Comin and Nanda (2019) analyzed the diffusion of 16 major technologies in 17 countries from 1870-2000. They found that greater financial market depth (i.e. more developed stock and credit markets) leads to faster adoption of capital-intensive technologies, but mainly in the early stages of diffusion. In later stages or for late-adopting countries, the effect fades. This suggests that robust financial markets are critical for financing the initial commercialization of new technologies. Alam et al. (2021) document a positive long-run impact of stock market development on the adoption of clean energy technologies across the EU and G20. Firm-level evidence also links investment capacity to technology adoption: Lanteri and Rampini (2025) develop a model of clean equipment adoption under financing frictions, supported by empirical data on commercial shipping fleets. They document that financially constrained companies adopt cleaner, energy-efficient technology more slowly. Higher stock valuations, which would enable equity issuance, can accelerate corporate adoption of clean technologies.

2.1.4 Factors Driving Investment in Clean Technology

As mentioned, global clean energy investment worldwide grew to $2.1 trillion in 2024, compared to $231 billion in 2010 (Statista, 2025a), which reflects the convergence of supporting policies, technological progress and shifting investor attention towards sustainability. However, not all clean technologies are equal in the eyes of investors. Since 2000, solar power and wind power have consistently drawn the largest share of energy investment, while areas like biofuels, hydrogen, or carbon capture have lagged behind. Several factors explain the factors driving investment in clean technology and the divergence between different clean technologies.

Policy and Regulation

Government action has been a fundamental driver of clean technology investment. Early milestones like the Kyoto Protocol (1997) and national policies in the 2000s created demand signals for renewables (L. C. Hunt et al., 2024). Over time, global accords, such as the Paris Agreement in 2015, and local targets (e.g. the EU Green Deal, China's five-year plans) have solidified investor confidence that clean technology markets will keep expanding (European Commission, 2023; Hepburn et al., 2021; UNFCCC, n.d.) Kruse et al. (2024) showed that companies with substantial green revenues saw their stock price jump by roughly 10% in cumulative abnormal returns after the 2015 Paris Agreement's signing. This suggests that investors reallocated capital in expectation of clean technology expansion, effectively rewarding companies positioned in green markets. In recent years, unprecedented policies have encouraged the sector's expansion. The U.S. Inflation Reduction Act (2022) alone allocates an estimated $1.2 trillion to climate-related investments – the largest climate investment in U.S. history – spanning energy, transport and industry (Goldman Sachs, 2023).

Conversely, lack of consistent policy can decrease investment. Early venture investors learned during the mid-2000s 'Cleantech 1.0' boom and bust that policy risk – the chance that supportive regulations won't materialize or persist – can derail demand for clean solutions (van den Heuvel & Popp, 2022). Over time, however, governments worldwide have generally intensified climate policies only upward, giving investors more certainty. The overall trend since 2000 is a strengthening policy landscape favoring clean technology, despite some short-term reversals, which has been fundamental in driving capital into the sector (L. C. Hunt et al., 2024).

The strength and predictability of policy support differs by technology, influencing investor preferences. For instance, in the U.S. solar energy has been consistently receiving policy support since the 1970s with the Solar Energy Research, Development and Demonstration Act. (United States Congress, 1974). On the other hand, emerging technologies have received only recent or temporary policy support. For example, only in the past few years have governments announced major hydrogen strategies and subsidies. This support, while significant, is still viewed as uncertain and in early stages – investors are waiting to see if promised incentives (e.g. hydrogen tax credits, electrolyzer subsidies) are implemented fully and sustained beyond initial periods (Reuters, 2024)

Technology Maturity and Scalability

Another driver is the technology maturity and scalability of different clean technologies. According to the United Nations' Intergovernmental Panel on Climate Change (IPCC), from 2010 to 2019 the unit costs of solar PV fell around 85%, onshore wind by ~55%, and lithium-ion batteries by ~85% (IPCC, 2022). This sustained cost decline has made technologies like solar, wind, and battery storage increasingly competitive without subsidies. In 2023, each dollar invested in wind and solar today produces 2.5 times more energy than a decade ago, underscoring the efficiency gains achieved (International Energy Agency, 2025).

In contrast, emerging technologies such as hydrogen fuel cells and electrolyzers lack extensive operational track records. Effective transition to a hydrogen economy requires policy, infrastructure and financial support mechanisms, including incentives and subsidies, to mitigate these risks and encourage private sector investment (Nyangon & Darekar, 2024). Hydrogen also faces infrastructure challenges: production (electrolyzers), distribution (pipelines, storage), and end-use equipment (fuel cells, industrial furnaces) must all scale together. This makes the investment case more complex and uncertain – even if one part of the value chain is ready, the others might lag, stalling the whole system. The International Energy Agency observes that most hydrogen projects remain at early stages and face financing problems (Reuters, 2024). Some clean technology sub-sectors inherently target smaller or more fragmented markets. Geothermal energy, for instance, requires specific geological conditions, which are not universally available and therefore restrict the global scalability of geothermal projects (IFP, 2023).

These factors have resulted in a 'two-speed' clean energy transition in investment (International Energy Agency, 2025). Mature technologies that are affordable, proven, and scalable have captured the vast majority of funding, whereas emerging solutions still struggle to attract private investment at scale. Mature sectors – renewables, energy storage, EVs, and grids – drew \$1.93 trillion of investment in 2024 (over 90% of total), while emerging sectors – including hydrogen, CCS, nuclear, and others – collected only about \$155 billion (≈ 7%). For instance, for hydrogen, the question arises who will buy it and at what price. In many proposed uses (e.g. replacing natural gas in industry), the end customers are not yet willing to pay a green premium, making hydrogen project revenues more speculative (O. B. Hunt & Tilsted, 2024).

Risk-Return Profile and Revenue Streams

Investors favor clean technology businesses with proven business models and revenue streams. In the early 2000s, clean technology was viewed as a high-risk investment, as clean energy projects often had long development timelines and capital-intensive scale-up, meaning it was hard to achieve explosive growth (van den Heuvel & Popp, 2022). As clean technologies became more cost-competitive and widespread, their cash flows became more predictable, improving the risk-return ratio. For instance, energy storage is interesting to investors now because batteries can earn revenue in multiple established ways (wholesale market arbitrage, grid services, capacity payments) (McKinsey & Company, 2025), but in 2010, storage investment was low (Loock, 2012).

ESG Pressure and Sustainable Investment

The growth of environmental, social, and governance (ESG) investing has been a major factor driving investment into clean technology. Since the early 2000s, institutional investors have increasingly committed to considering sustainability and climate risks in their portfolios. By 2020, global assets under management in sustainable investment strategies reached $35 trillion (World Economic Forum, 2022). This wave of capital is partly motivated by values and partly by recognition that climate change poses financial risks. Clean technology companies and projects often score well on ESG criteria (by virtue of enabling decarbonization), making them prime targets for institutional investors, including funds and sovereign wealth funds. This ESG momentum translates into real investment: many pension and sovereign wealth funds now have dedicated allocations for renewable infrastructure or climate-tech venture, driven by stakeholder pressure to address climate change. Investors also see long-term opportunity in aligning with climate goals (IFSWF, 2020; OECD, 2011).

Media Attention and Public Sentiment

Public awareness and activism around climate change have put additional pressure on investors to fund cleaner technologies. This heightened media coverage has increased public awareness and in turn pushed investors and politicians to act, translating into public demand for solutions – and willingness to invest in them. Retail investors, influenced by this attention, have invested into clean technology stocks and funds. Research finds that increased climate news coverage leads to significant inflows into green mutual funds (Cornelli et al., 2025). Media and broader public sentiment also indirectly drive investment by influencing consumer adoption. Positive press about solar energy, electric vehicles, or heat pumps helps educate consumers and build demand for these products. Strong demand signals then encourage investors to back companies in those spaces. For instance, years of enthusiastic media coverage of electric vehicles and the climate benefits of going electric have helped EVs achieve mainstream acceptance. This contributed to exponential growth in EV sales – from almost zero a decade ago to over 10 million electric vehicles sold in 2022 (Rodrigue, 2024).

While structural drivers of policy framework, technological maturity and economic incentive explain some variation in the adoption and investment of clean technologies, they do not capture the behavioral and informational dynamic guiding capital allocation. As emerging clean technologies compete for finite amounts of investor capital in a crowded and fast-moving field, media visibility and framing of technologies become important factors in shaping investor decision making. The public talk and exposure from the media could either magnify or tone down perceived opportunities and influence which technologies would come into the investor choice set. The next part discusses how media coverage shapes investor attention, how this attention affects trading behavior, and why these considerations are particularly important for clean technology companies at the intersection of policy, innovation, and public perception.

2.2 Investor Attention to News Coverage

The literature demonstrates that news coverage, a proxy for investor attention (Cumming et al., 2016), significantly influences stock market behavior, often leading to increased trading volumes and temporary price increases, consequently encouraging innovation adoption (Comin & Nanda, 2019; Alam et al., 2021).

2.2.1 Behavioral Finance Perspectives on Investor Attention

Investor attention refers to the limited attention of investors when having to choose between particular events, companies, stocks or industries (Maula & Lukkarinen, 2022). Barber and Odean (2008) highlight in the Handbook of Economics and Finance (Constantinides et al., 2013) that individuals have a finite capacity for attention when it comes to making investment decisions. This phenomenon of attention as a scarce cognitive resource, generalized by Kahneman (1973), influences trading behaviors in two main ways. On one side, insufficient attention to crucial information can cause a lag in response to significant news. Conversely, allocating too much attention to potentially outdated or irrelevant details can prompt an exaggerated reaction. They argue that attention significantly impacts the buying decisions of individual investors: instead of conducting a systematic search, many investors tend to focus on stocks that immediately capture their attention and stand out (Barber and Odean, 2008) – such as those frequently mentioned in the news or experiencing significant price fluctuations. This tendency leads individual investors to predominantly purchase stocks that stand out. This limited attention capacity also leads investors to allocate more attention to sector-level factors (such as industries, locations or countries) than to firm-specific factors (Peng & Xiong, 2006). According to Merton's Investor Recognition Hypothesis (1987), assets followed by a larger set of investors (i.e. higher investor recognition or awareness) will enjoy greater demand and liquidity, resulting in a lower expected return (cost of capital) as investors are willing to accept lower returns for well-known stocks. In other words, stocks that are neglected must offer higher returns to compensate for their limited visibility. This theory highlights a simple but powerful mechanism: when a firm attracts more investor attention, it can potentially raise capital more easily and at better valuations due to a broadened investor base.

Given that investor attention is limited and selective, the role of financial media becomes central in determining which technologies receive capital allocation. In the next sections, media coverage is examined as a primary channel through which information and ideas gain the attention of investors. The literature has explored how financial media influence markets by shaping what information investors see and how they interpret it.

2.2.2 The Role of Media Coverage in Directing Investor Attention

Given the scarcity and selectivity of investor attention, financial media play a crucial role in shaping which companies and sectors attract investor focus. News Coverage can be used as a proxy for investor attention (Cumming et al., 2016) and can significantly impact trading volume and asset prices, sometimes more than the underlying informational content itself (Engelberg and Parsons, 2011)

Barber and Odean (2008) find that individual investors buy attention-grabbing stocks, such as stocks in the news, and that preferences come into play only after attention has determined the choice set. Engelberg and Parsons (2011) build on this by establishing a causal link between news coverage by local media and stock market reactions. The researchers found a causal link between local media coverage and trading activity, independent of the investors' proximity to the traded companies or the companies' local market significance. The study also highlights that while extreme earnings reports influence trading volumes, the impact of media coverage is significantly more pronounced, indicating that media exposure can drive trading behavior more powerfully than the information itself. Da et al. (2009) use the Google search Volume Index (SVI) to show that SVI captures the attention of retail investors. An increase in SVI predicts higher stock prices in the next two weeks and an eventual price reversal within the year, representing a temporary raise of stock prices. This suggests that media-driven investor attention results in temporary price pressures, consistent with attention-based trading.

However, the literature presents countervailing evidence. Fang and Peress (2009) find that stocks with very little or no media coverage earn higher returns – 3% annually – than stocks with high media coverage, even after controlling for risk factors, such as market, size, momentum and liquidity. This suggests that high media exposure may inflate stock prices through investor overreaction, leading to lower future returns. Yet, the authors also emphasize a positive informational role of the media, particularly for companies lacking analyst coverage, where media attention serves as a substitute information channel. Extending their research to institutional investors, Fang et al. (2014) find that professional funds tend to buy stocks with media coverage more heavily than those without. However, funds who buy a lot of media-covered stocks underperform compared to other funds that don’t exhibit the same behavior.

These studies underscore the dual role of media in financial markets. On the one hand, media coverage directs investor attention, thereby increasing trading activity and accelerating market reactions. On the other hand, excessive reliance on media exposure can introduce noise into markets, as investors may overreact to the presence of coverage rather than the informational value of the news itself. The next point elaborates on the effect of the content and characteristics of media coverage, namely its sentiment and political orientation, in directing investor behavior.

Political Bias and Sentiment in Media Coverage

The sentiment or tone of news reports can influence investor sentiment about a certain topic. Media reports often contain qualitative assessments – optimism or pessimism about a firm or industry – that go beyond factual data. These sentiment cues can influence investor psychology and expectations. Tetlock (2007) finds that high media pessimism predicts a decrease of market prices, followed by a reversion to previous levels. Unusually high or low pessimism predicts high market trading volume. Data also shows that the market reactions to negative sentiment are not concentrated post-information release, but are distributed throughout the trading day. Negative returns following pessimistic content typically reverse within a few days, undermining the idea that such sentiment reflects fundamental information.

Media bias exists, to a greater or lesser extent depending on the specific region. Newspapers can exhibit political or ideological biases in how they report on certain topics. For instance, in the U.S., debates often have a political dimension (e.g. disagreements over policy, government subsidies), and media outlets sometimes frame technology innovations through partisan lenses. Gentzkow and Shapiro (2010), for instance, construct an index of media slant and found that news outlets cater to the ideological leanings of their audience. There is significant demand for slanted news, and newspapers adjust their language to align with either a more conservative or liberal audience as a profit strategy. Bias can appear in multiple ways: selection bias determines which stories the newspaper selects to publish, coverage bias refers to how much weight is given to a story, and statement bias determines how a story is reported. Saez-Trumper et al. (2013) find that political bias is most visible on social media. For traditional media, bias is especially observable in terms of coverage and statement bias. The coverage bias is seen when comparing the distribution of length of articles on different stories, while the statement bias is especially visible in the sentiment polarity with which U.S. presidency candidates are treated.

The literature establishes a link between media coverage and investor attention, highlighting how news visibility, sentiment tone, and political inclination affect trading behavior and stock prices. Media do not merely transmit information; they shape which topics investors consider, how they interpret them, and how they respond in financial markets. While investor attention can improve capital allocation efficiency, excessive or biased media coverage may also distort prices through temporary overreactions. This theoretical foundation provides the basis for investigating how media coverage of clean technologies influences the investment behavior of market participants. The next section builds on these insights to formulate hypotheses linking clean technology media attention, sentiment, and political orientation to abnormal stock returns.

3. Hypothesis Development

Hypothesis 1: Investor Attention Effect

The high uncertainty characterizing the clean technology sector, makes it susceptible to external signals, especially those shaping investor expectations (Kemp and Volpi, 2008) and by extension, access to capital. Among these signals, media coverage plays a central, but underexplored role.

Mass media serve as an important mechanism for framing innovations and signaling legitimacy to a broader set of investors. As Rogers (2003) notes, mass media channels are more effective at diffusing simpler innovations, whereas interpersonal networks are needed for complex ones. An important distinction needs to be made between different types of mass media. While social media is focused on quick dispersion of key ideas, legal newspapers have the ability to transmit complex ideas in a digestible shape. Therefore, legal media coverage of clean technology companies and trends may function as a critical bridge between industry and investor.

While prior literature has established that media attention influences stock market behavior across various sectors (Barber & Odean, 2008; Da et al., 2009; Engelberg & Parsons, 2011), the specific role of mainstream, legal newspaper coverage in shaping investor attention towards clean technology companies remains insufficiently examined. Most existing studies rely on proxies like Google search volumes or social media data (e.g. Liu et al., 2021; Pham & Huynh, 2020, Liu et al., 2024) or analyze legal media coverage in relation to VC investment in clean technology (Cumming et al., 2016), with relatively few investigating how legal news sources affect the stock market valuation and visibility of companies in emerging clean technology sectors. Yet, such coverage likely plays a more formative role in more complex sectors where policy, science and public perception intersect. It can help simplify narratives, reduce informational asymmetries, and make these clean technology companies more known to investors.

From a clean technology adoption perspective, this matters. Access to financial capital is a known constraint for emerging and capital-intensive technologies (Comin & Nanda, 2019; Lanteri & Rampini, 2025). If investor attention triggered by media coverage can meaningfully influence stock returns, it implies that visibility itself becomes a lever for accelerating adoption by easing financing constraints, increasing firm valuation, and signaling market readiness. Understanding this mechanism is thus essential for policymakers, entrepreneurs, and investors seeking to scale climate-relevant emerging technologies.

Based on this evidence, it is expected that clean technology newspaper coverage, by attracting attention to clean technology developments, should influence stock returns of clean technology companies. This leads to the formulation of the first hypothesis:

<u>H1 – Investor Attention Effect</u>: Clean technology newspaper coverage influences stock market returns of clean tech companies.

This hypothesis advances the literature by connecting investor attention theory to the specific context of clean technology finance, where uncertainty and innovation intersect. It is testable using event study methodology, examining whether abnormal stock returns follow clean tech-related media events. If confirmed, this would support the notion that clean technology companies can gain market momentum not only through fundamentals or subsidies, but also through strategic visibility in trusted legal media channels, ultimately influencing clean technology adoption.

Hypothesis 2: Tone Effect

While visibility through media coverage is necessary to attract investor attention, the question of whether all publicity is good publicity, can be raised. It is not enough to simply assess whether news coverage has an influence on market returns of clean tech companies (Hypothesis 1); the tone of the news article on clean technology (positive, neutral or negative) could affect investor sentiment, perceived risk, and expectations about future performance. In highly uncertain sectors like clean technology, the tone of media coverage may carry disproportionate influence. The financial literature recognizes that media tone impacts trading behavior. Tetlock (2007) finds that negative media sentiment induces short-term downward price pressure and higher trading volume, even when fundamentals do not change. Applied to clean technology, this suggests that positive tone may enhance investor confidence and returns, while negative tone may cause short-term negative stock price effects and higher trading volume.

However, despite this growing recognition, few studies apply sentiment analysis to legal, mainstream news coverage of clean technology companies. Most prior work has focused on social media (Chen & Mai, 2024) or has analyzed sentiment in relation to policy statements (Noailly et al., 2024). As a result, the role of legal media sentiment in shaping investor behavior in the context of clean innovation remains underexplored. Positive tone may build confidence in the viability and policy support of clean technologies, lowering perceived risk and increasing investment. Conversely, negative tone may amplify uncertainty. From an innovation diffusion standpoint, tone may also shape relative advantage and perceived compatibility – two important drivers of adoption (Rogers, 2003).

This motivates the second hypothesis:

<u>H2 – Tone Effect</u>: The tone of newspaper articles on clean technologies influences the stock market returns of clean tech companies.

This hypothesis aims to fill a gap in the literature by combining sentiment analysis with legal media, specifically in the context of clean technology. It is testable through a panel regression framework. If confirmed, it would show that not only visibility, but how clean technology is portrayed in legal media channels influences market outcomes and, by extension, shapes the financial landscape for clean technology diffusion.

Hypothesis 3: Political Alignment Effect

In addition to the tone, the political inclination of the media outlet publishing the article may further determine how investors interpret and respond to clean technology news. In politically polarized environments, particularly in countries like the U.S., clean technology is not just a technical or economic issue, but also an ideological one. As a result, the framing of clean technology news, and the subsequent investor response, may vary systematically depending on the political alignment of the news outlet.

News outlets do not simply transmit any information, they select, prioritize and frame content according to editorial lines that reflect the company culture and the preferences of their target audience (Gentzkow and Shapiro, 2010). This introduces different types of media bias: news outlets may differ in what stories they choose to publish (i.e. selection bias) and in how they frame clean technology news (i.e. statement bias) (Saez-Trumper et al., 2013). Media that aligns with conservative political ideologies may frame clean energy more skeptically or highlight cost concerns, whereas progressive-leaning outlets may emphasize environmental urgency and innovation potential. To give an example: in the beginning of 2025, multiple newspapers reported on the Biden administration's new climate rules that added in exemptions for hydrogen. Left news sources used positive language, the center was favorable, but insisted for cautious support, and the right news sources identified potential loopholes for 'dirty hydrogen' (Ground News, n.d.). These differences in framing may in turn influence how investors interpret the same underlying news event depending on the source.

No empirical work has examined how the political alignment of the media source influences financial market reactions to clean technology news. This represents a meaningful gap in both financial and innovation diffusion research. If investors continuously rely on the same politically aligned news to interpret clean technology developments, then the media outlet's political orientation could condition the magnitude and direction of investment or adoption. Politicization of clean technology risks creating uneven capital access, where certain technologies or companies may receive more favorable market responses simply because they are framed more positively by ideologically aligned media. In turn, this could distort the allocation of investment.

Furthermore, from a policy standpoint, understanding how political media alignment shapes investor behavior is crucial. It can inform how governments and clean technology advocates design communication strategies to build support for green technologies.

This leads to the third hypothesis:

H3 – Political Alignment Effect: The political alignment of the publishing newspaper influences the stock market returns of clean tech companies.

This hypothesis is empirically testable through a panel regression framework. If confirmed, it would reveal that political media bias is not just a sociopolitical concern, but also an economic mechanism that may amplify or suppress the diffusion of clean technology through its effect on investor attention and capital flows.

Hypothesis 4: Interaction Effect

While both media tone and political alignment independently influence how investors interpret news, their interaction may shape investor behavior in more complex and context-specific ways, and may reinforce or contradict their individual effects. Despite growing attention to tone and political inclination in media research related to clean technology, most existing literature isolates one variable, sentiment (e.g. Tetlock, 2007; Chen & Mai, 2024; Barber et al.) or media slant (Gentzkow & Shapiro, 2010), without analyzing how they jointly condition investor responses.

The interaction effect matters because investors are not only influenced by what is said, but by who says it and how consistent the message is with their expectations. A positive article in a left-leaning outlet, for instance, may be perceived as credible reinforcement by investors already inclined to view clean technology favorably. In contrast, the same positive story in a right-leaning outlet may trigger a stronger market reaction, because it represents a surprising signal: a positive tone emerging from a source not typically aligned with pro-clean-technology narratives. Conversely, negatively toned articles from left-leaning outlets may raise red flags, while similar negativity from conservative sources may be discounted. From the perspective of clean technology adoption, understanding this interaction is critical. If tone-politics interactions systematically shape market valuation, then the same clean technology development may receive very different investor reactions depending on how and where it is covered. This has direct implications for access to capital and, in turn, for the amount and speed of clean technology diffusion.

This leads to the formulation of the fourth hypothesis:

<u>H4 – Interaction Effect:</u> The interaction between the tone of the article and political alignment of the publishing newspaper influences the stock market returns of clean tech companies.

The hypothesis is testable through interaction terms in a panel regression model. If supported, this would provide evidence that the alignment between news article tone and newspaper political inclination can amplify or dampen the market reaction to clean technology news.

Hypothesis 5: Matching Technology Effect

While aggregate media attention to clean technologies can influence investor sentiment and market behavior, not all clean technologies are equal in terms of maturity, scalability, or investment appeal, as explained in subsection 2.1.4. This raises the question if the stock market reacts more strongly when media coverage aligns directly with the core technology of the firm, or if general clean technology visibility is sufficient to boost the entire sector.

This distinction matters because clean technologies exist in different stages of maturity. Some, like solar, wind, and electric vehicles are mature and widely adopted, while others, such as e-fuels or hydrogen, are still emerging and face technical, financial, and infrastructural barriers. Comin and Nanda (2019) find that financial market depth plays a critical role in accelerating the early-stage diffusion of capital-intensive technologies, but that this effect diminishes in later stages of adoption. Their findings imply that targeted visibility and investment are particularly important for emerging technologies where commercialization is still fragile and market validation is ongoing. Despite this, most existing research treats clean technology as a single category, often overlooking differences across technological subgroups. Prior studies on investor attention and media sentiment (e.g., Cumming et al., 2016; Liu et al., 2021; Noailly et al., 2024) focus on sector-level reactions. This creates a gap in understanding the mechanisms through which investor attention contributes to the diffusion of newer technologies versus reinforcing momentum in already-mature segments.

From a behavioral finance perspective, investors may be more responsive when there is a direct match between media content and the firm's core technology, due to greater perceived relevance and informational value. A hydrogen-focused investor, for example, is more likely to react to news about hydrogen policy or projects than to general commentary about solar or electric vehicles. From a clean technology adoption standpoint, if only matched coverage produces significant market reactions, it implies that emerging technologies must compete for targeted attention to access the capital needed. On the other hand, if general visibility benefits all clean technology companies equally, this may suggest a general spill-over effect that facilitates broader diffusion. Understanding which dynamic dominates has implications for media strategy and policymaker communication.

This leads to the formulation of the last hypothesis:

H5 – Matching Technology Effect: The influence of newspaper coverage on stock returns is stronger when the technology discussed in the article matches the core technology of the clean tech company.

This hypothesis extends previous attention-based research by introducing a technology-specific lens to the analysis of media effects. It is testable by using the same methods used in the previous four hypotheses. If supported, this would highlight that the effectiveness of investor attention as a channel for clean technology adoption depends on the precision of the signal, which has implications for how emerging technologies position themselves in the media landscape to attract investment.

Together, these five hypotheses aim to capture the multifaceted relationship between clean technology media coverage and investor behavior, accounting for attention dynamics, sentiment effects, political framing, their interaction effect and technology-specific sensitivities.

4. Research Method & Empirical Analysis

4.1 Data Collection

This report uses five types of data:

- CompanyData: The sample list of 108 companies compiled via S&P Capital IQ (Verdonck, 2025e) (4.1.1.).
- StockData (2006-2022): The stock market closing prices of the sample list of companies, from the Bloomberg Terminal (Verdonck, 2025b) (4.1.2.).
- SPXData (2006-2022): S&P500 market price data, from the Bloomberg Terminal (Verdonck, 2025c) (4.1.2).
- RFRData (2006-2022): The risk-free-rate, from the Federal Reserve Bank of St. Louis (Verdonck, 2025d) (4.1.2).
- NewspaperData (2007-2022): The newspaper data of U.S. and UK news headlines (Verdonck, 2025a) (4.1.3).
- Additional data for control variables in panel regressions (2006-2022):
  - Yearly total assets of each company in CompanyData, from the Bloomberg Terminal under the name BS_TOT_ASSETS (Verdonck, 2025f). The total assets are reported in millions of dollars.
  - Yearly total employees of each company in CompanyData, from the Bloomberg Terminal under the name NUM_OF_EMPLOYEES (Verdonck, 2025g).

For full variable descriptions, sample extracts and data sourcing, see Appendices 5-11.

4.1.1 Sample Company List: CompanyData

S&P Capital IQ was used to determine the sample list of companies for which stock data was then collected. The screening criteria applied are detailed in Appendix 1. This resulted in a list of about 130 companies. After thorough research into each of them individually to assure their presence in the clean technology sector, a list of 114 companies was compiled. The complete list of companies can be found under Appendix 2.

These companies are categorized by industry: E-fuels, Energy Storage (i.e. batteries and its components), Energy Technologies (i.e. clean technology IT platforms, clean lighting, generators, semiconductors, etc.), Electric Vehicle Manufacturers and Technologies, Renewable Energy, Solar Energy, Wind Energy, and Marine Transportation. An additional set of five binary variables was compiled, which indicate the companies' technology focus across different industries: *CompanyHydrogen, CompanyAmmonia*, *CompanyMethanol, CompanyEV*, *CompanyEfuel* (includes companies in multiple e-fuel industries, such as hydrogen, biofuel, ammonia and methanol).

4.1.2 Financial Stock Data Collection: StockData, SPXData, RFRData

The financial data was sourced from the Bloomberg Terminal. For each of the 114 companies, daily closing price data from 2006-2022 was collected, resulting in a dataset of 5027 rows.

Six of these companies were removed from the dataset. GE Vernova Inc, Fly-E Group Inc., NANO Nuclear Energy Inc., Solarmax Technology Inc., Nextracker Inc. were removed due to insufficient observations prior to December 2022 – the latest available news article publication date. When computing the CAR in subsequent analysis, SKYX Platforms Corp. appeared to have extraordinarily small and large values, so to maintain robustness in further analysis, that company was also removed, resulting in a final sample of 108 companies. These six removed companies are marked in blue in the table under Appendix 2.

The market returns were computed using the S&P 500 Index, the data for which was extracted from the Bloomberg Terminal. The risks-free rate is the 10-year constant maturity yield on U.S. Treasury Securities, the data for which was extracted from the Federal Reserve Bank of St. Louis's website. Both datasets cover a period from 2006-2022

4.1.3 News Article Data: NewspaperData

The original newspaper dataset contains 4.5 million headlines from 2007 until 2022 from ten different newspapers (Kaggle, 2023), which were selected based on the 2022 version of Statista's *Leading global English-language news websites in the United States in 2024, by monthly visits* (Statista, 2025c). Since 2022, the order of the ranking has changed slightly, but all newspapers, except for the Guardian, still remain in the top twenty.

The decision to focus on the U.S. as the geographical scope of this analysis is supported by several economic, institutional, and informational factors. Economically, the U.S. emerged as a global leader in clean technology investment, particularly following the passage of the Inflation Reduction Act in 2022 (Goldman Sachs, 2023). Institutionally, the U.S. capital markets are among the most liquid and data-rich globally (Euronext, 2024). From an informational perspective, the U.S. media landscape is not only expansive but also influential beyond national borders; a survey across 10 European countries revealed that, on average, 51% of respondents regularly follow U.S. news (Poushter, 2018). Furthermore, unlike many European outlets that prioritize editorial neutrality, U.S. newspapers often display clear political leanings. To account for this variation, the Ground News bias ratings – based on assessments by All Sides, Ad Fontes Media, and Media Bias Fact Check – were employed to systematically capture the ideological orientation of news sources included in the analysis (Ground News, n.d.).

Table 1 represents the ten newspapers ranked in popularity (Statista, 2025c), their country of residence (U.S./UK) and their political inclination (Ground News, n.d.).

Table 1: Newspaper Ranking

Ranking	**Newspaper**	**Monthly website visits in the U.S. in 2024 (in millions)**	**U.S./UK**	**Political Inclination**
1	The New York Times	463.07	U.S.	Left
2	CNN	356.62	U.S.	Left
3	FOX News	253.58	U.S.	Far right
4	USA Today	143.62	U.S.	Left
5	New York Post	142.11	U.S.	Right
6	BBC	111.24	UK	Center
7	Daily Mail	108.63	UK	Right
8	CNBC	83.82	U.S.	Center
9	Washington Post	82.42	U.S.	Left
10	The Guardian	N/A	UK	Left

Source: Author's own elaboration; using data from Statista, 2025c; Ground News, n.d.

In order to filter the data for news articles on the topic of clean technology, text analysis was performed in Python – the code can be found under Appendix 4, only keeping the articles that contain one or more of these keywords in their headline:

clean tech, clean technology, green technology, renewable energy, sustainable energy, green power, power-to-x, solar power, solar panel, wind power, wind turbine, wind farm, wind mill, hydropower, bioenergy, biofuel, geothermal power, electric vehicle, EV charging, EV, battery storage, energy storage, lithium battery, smart grid, power-to-gas, power-to-liquid, hydrogen, methanol, ammonia, hydroelectric, alternative fuels, fuel cells, microgrid, carbon capture, energy efficiency, low-carbon technology, sustainable transport, renewable resources, waste-to-energy, wave energy, biogas, perovskite solar cells, net zero emissions, smart energy systems, sustainable development, energy transition, green building, smart thermostats, energy management systems.

The text analysis together with a removal of double headlines, reduced the dataset from around 4.5 million rows to 1558 rows.

Afterwards, sentiment text analysis was performed on the headlines, resulting in a *positive*, *negative* or *neutral* sentiment identification (i.e. *Sentiment* variable). Sentiment analysis analyzes opinions, sentiments, evaluations, attitudes and emotions via the computational treatment of subjectivity in text (Hutto & Gilbert, 2014). For the text analysis, the VADER (Valence Aware Dictionary and sEntiment Reasoner) model was used. This lexicon-based tool is particularly well suited for texts from social media, news headlines and texts where sentiments are expressed concisely. It uses a combination of sentiment lexicon and grammatical rules and it is sensitive to both polarity (positive or negative) and intensity of emotion. VADER assigns a sentiment score to each word based on its emotional value. It then combines these scores into an overall compound sentiment which ranges from -1 to 1. If the compound score is higher than 0.05, the sentiment is positive. If it is lower than -0.05, the sentiment is negative. If the compound score is in between -0.05 and 0.05, the sentiment is neutral.

Compared to machine-learning text analysis approaches, such as the Naïve Bayes, which requires extensive training data and are computationally expensive, the VADER approach is more streamlined and lends itself well for analysis of short texts (Chadha & Chaudhary, 2023). Hutto and Gilbert (2014) analyze VADER's effectiveness compared to eleven other text analysis methods in context of social media analysis and find that VADER often outperforms the other eleven text analysis methods when classifying the sentiment of tweets into positive, neutral of negative classes. Since headlines are of similar conciseness as tweets, this model is well suited for the analysis in this report.

Using Ground News' website and its political inclination identification, a *Politics* variable was added. The variable separates the news articles into '*left*', '*center*', '*right*' and '*far right*' inclination. Ground News is a platform that aggregates news stories to counteract media bias by offering users a comprehensive view of how different news sources cover the same story. It aggregates news from a wide range of sources across the political spectrum, presenting them side-by-side to highlight differences in reporting. The platform allows users to see how various news outlets report on the same event, which can help them understand potential biases and gain a more balanced perspective.

It also identifies stories that are being reported by media outlets on one side of the political spectrum but are underreported or ignored by the other side.

The third binary variable *NewsCountry* identifies the origin of the newspaper. It is '1' if the newspaper is American, i.e. The New York Times, CNN, Fox News, USA Today, The New York Post, CNBC and The Washington Post. In contrast, it is 0 if the newspaper is from the UK, i.e. The Guardian, The Daily Mail, The BBC.

The last five variables indicate which type of clean technology is mentioned in the headline and corresponds to the same technologies as in the CompanyData dataset: *NewsHydrogen, NewsAmmonia, NewsMethanol*, *NewsEV*, *NewsEfuel* (includes headlines with the keywords 'e-fuel', 'alternative fuel', 'biofuel', 'ammonia' and 'methanol')

4.2 Methodological Approach

The empirical analysis is conducted using the statistical program R – the code can be found under Appendix 3 – which enables robust data handling and econometric modeling. Following established methodological frameworks (Fama et al., 1969; Henderson, 1990), the analysis proceeds in two main stages: exploratory data analysis and hypothesis-driven empirical testing.

4.2.1 Exploratory Data Analysis

Initially, exploratory analysis is conducted to examine the distributional characteristics of both the company- and newspaper datasets. This phase includes the verification of key variables, such as company technology classification, news article tone and political alignment.

4.2.2 Hypothesis Testing

PART 1: Influence of Media Coverage on Stock Returns

PART 1.1: Event Study – Significance of the Mean CAR

The first part employs an event study to test whether investor attention to clean technology news coverage influences stock returns. The event study methodology examines the stock price reaction around the announcement of an event (Flammer, 2021) and seeks to measure the return impact on stocks after that announcement (Pettengill & Clark, 2001). Event study is a widely accepted method in financial research for identifying the impact of discrete informational events (e.g., clean technology news headlines) on asset prices (MacKinlay, 1997; Henderson, 1990).

The main hypothesis is as follows:

H1 – Investor Attention Effect: Clean technology newspaper coverage influences stock market returns of clean tech companies.

Clean technology newspaper coverage serves as a proxy for investor attention. This proxy is motivated by literature suggesting that media coverage plays a central role in directing limited investor attention, particularly in sectors characterized by high uncertainty and information asymmetries. Frequent reporting on cleantech developments signals to investors that the sector is gaining importance and relevance (e.g., Cumming et al., 2016; Ming et al., 2023).

To quantify the market reaction to clean technology news, abnormal returns (AR) are computed for each firm i around the publication of news article n. In order to assure the robustness of the analysis, two methods for computing the abnormal returns are compared against each other: the general market model (Flammer, 2021) and the Capital Asset Pricing Model (CAPM) (Pettengill & Clark, 2001).

In the market model, the coefficients α_i and β_i are estimated by ordinary least squares (OLS) based on the estimation window of before the event window. This estimation is repeated for every article n. This equation is formally estimated:

$$R_{it} = \alpha_i + \beta_i \times R_{mt} + \varepsilon_{it} ,$$

Where

- R_{it} it represents the return on the stock price of company *i* on day *t*,
- α_i is the intercept term, representing the average return on the stock of company *i* that is unexplained by market movements,
- β_i reflects the sensitivity of the stock of company *i* to market movements,
- R_{mt} is the daily market return,
- ε_{it} is the residual.

The estimated return on the stock of firm *i* on day *t* is then given by

$$\hat{R}_{it} = \hat{\alpha}_i + \hat{\beta}_i \times R_{mt} .$$

Then, the abnormal daily return (AR) of firm *i* on day *t* is computed as follows:

$$AR_{it} = R_{it} - \hat{R}_{it} .$$

Lastly, the CARs for each time interval are computed by summing up the abnormal returns within a specific event window per event date of article n.

However, Pettengill & Clark (2001) assess this market model method as inherently biased. Specifically, it highlights the issue of non-stationarity in OLS parameter estimates, as discussed by Hays and Upton (1986), with an emphasis on the systematic bias introduced in the intercept (α), a factor often overlooked relative to β instability (Binder, 1998). When the market model is used to estimate α and β in such contexts, the resulting α tends to be upwardly biased, reflecting returns not justified by systematic risk exposure (i.e., β). Consequently, when abnormal returns (AR) are calculated by subtracting the expected return (based on the biased α and β estimates) from actual returns, there is a downward bias in post-event CARs. While this may not distort the impact measured on the event day itself, the cumulative abnormal returns over a window can be materially affected by this bias.

The authors advocate for the Capital Asset Pricing Model (CAPM) instead of the simpler market model. The expected abnormal returns, following the CAPM, are computed as

$$AR_{it} = R_{it} - (R_{ft} + \hat{\beta}_i \times (R_{mt} - R_{ft})) ,$$

where

- AR_{it} is the daily abnormal return for the selected stock,
- R_{it} is the daily return for the stock,
- R_{ft} is based on the 10-year constant maturity yield on U.S. Treasury Securities on a yearly basis. It is transformed to the daily return.
- $\hat{\beta}_i$ is the same beta as estimated in the market model and reflects the sensitivity of the stock of company i to market movements,
- R_{mt} is the market return measured by the return on the S&P500 Index.

The CAPM method elaborates on the market model by including the risk-free rate in the equation. It estimates the expected return on an asset relative to the risk-free rate and the excess return of the market (market return – risk-free rate). This method omits the α term entirely, thereby eliminating the intercept bias. The expected return is determined purely by the risk-free rate and the risk premium (market return - risk-free rate), scaled by the asset's β. This adjustment provides a more theoretically grounded and statistically robust method for estimating abnormal returns. Therefore, this report focuses on the CAPM method to analyze the CARs over dynamic event windows, which allows for a comparison of CAR significance and size across the different subwindows.

T-tests are conducted on the mean CAR of each different subwindow. Per subwindow, the mean CAR is computed by averaging the CARs over all event dates (i.e. article publication dates) and all companies. A t-test is a statistic that can be used to determine whether the mean is significantly different from zero (Campbell et al., 1997). This allows for a detailed examination of the event's effect both immediately after its publication and over subsequent days. Specifically, results are marked as *** if significant at the 1% level (p-value < 0.01), ** at the 5% level (p-value < 0.05), and * at the 10% level (p-value < 0.1). These thresholds reflect decreasing degrees of confidence in rejecting the null hypothesis of the mean being zero. Coefficients without stars are considered not statistically significant at conventional levels. This notation and interpretation of significance is followed throughout this report.

PART 1.2: Panel Regression Analysis – Role of Tone and Political Alignment

The second stage of the analysis builds upon the event study and examines whether the content characteristics of clean technology news articles help explain the variation in the CARs of clean technology companies. In particular, this section tests the influence of (H2) the tone of the article, (H3) the political alignment of the news outlets, and (H4) their interaction, controlling for company-specific and temporal factors. As a reminder of the hypotheses tested in this section,

H2 – Tone Effect: The tone of newspaper articles on clean technologies influences the stock market returns of clean tech companies.

H3 – Political Alignment Effect: The political alignment of the publishing newspaper influences the stock market returns of clean tech companies.

H4 – Interaction Effect: The interaction between the tone of the article and political alignment of the publishing newspaper influences the stock market returns of clean tech companies.

Panel regression models are estimated to investigate whether media sentiment and political alignment explain heterogeneity in CARs across events and companies. A panel regression is a statistical method designed to analyze data with both cross-sectional and time-series dimensions, allowing to control for unobserved heterogeneity across entities (e.g. companies) or time periods. The fixed effects model (a.k.a. within estimator), applied to individual entities (i.e. companies, in this report) is used. It isolates within-entity variation by removing all time-invariant characteristics, which mitigates omitted variable bias from unobserved factors correlated with the independent variables (Wooldridge, 2010). This leads to the following panel regression model.

$$CAR\_0\_10_{i,n} = \sum_{k=1}^{2} \beta_{1k}\, Tone_{k,n} + \sum_{l=1}^{3} \beta_{2l}\, Politics_{l,n} + \sum_{k=1}^{2} \sum_{l=1}^{3} \beta_{3kl}\, (Tone_{k,n} \times Politics_{l,n}) + \gamma_{i,n} + \varepsilon_{i,n}$$

Where

- $CAR\_0\_10_{i,n}$ denotes the CAR for company *i* associated with news headline *n* over [0,10] days after the news article publication.
- The variables $Tone_{k,n}$ are dummy variables for the *k* non-neutral sentiment categories (positive, negative)
- The variables $Politics_{l,n}$ are dummy variables for the *l* non-centrist political alignments (left, right, far-right)
- $\Upsilon_{i,n}$ represents the control variables that help isolate within-firm variation over time
    - News control variable: *NewsCountry*
    - Company control variables: *USCompany*, *CompanyIndustry, Total_Assets_log, Total_Employees_log*
    - Time control variable: *Year*
- $\varepsilon_{i,n}$ is the idiosyncratic error term

The coefficients β_{1k} measure the average effect of each non-neutral tone (relative to neutral) when outlets are centrist, while β_{2l} measure the effect of each non-centrist political alignment (relative to centrist) under neutral tone. The interaction coefficients β_{3kl} quantify how the impact of tone varies across different political slants.

PART 2: Subgroup Comparison – Technology-Specific Sensitivities

While in PART 1, the complete dataset is used across all types of clean technology, PART 2 now explores the heterogeneous effects across clean technology subgroups, but repeats the analysis methodology of PART 1.1 and 1.2. The goal is to test whether any of the five clean technology subgroups (Hydrogen, Ammonia, Methanol, E-fuels, Electric Vehicles (EV)) is more sensitive to news coverage and its characteristics (i.e. tone and political orientation) than others. These five technologies are intrinsically interlinked through their common applicability to the decarbonization of the transportation sector. While electric vehicles represent a more mature and infrastructurally ready technology, the emerging technologies of hydrogen and its derivatives (ammonia, methanol, and e-fuels) are increasingly positioned as complementary or alternative solutions. This divergence implies that investor reactions to news coverage may vary considerably depending on the perceived feasibility, familiarity, and momentum of each technology.

As a reminder of Hypothesis 5,

H5 – Matching Technology Effect: The influence of newspaper coverage on stock returns is stronger when the technology discussed in the article matches the core technology of the clean tech company.

PART 2.1: Event Study – Significance of the Mean CAR by Matching Technology

This part replicates the event study methodology from PART 1.1, but separates companies into subgroups based on their core technology. The aim is to compare the mean CARs for companies that operate in the same technology covered by the article (Matched) and that operate in unrelated or substituting technologies (Not Matched).

Remember that a set of binary variables was created for:

- Five binary variables for the news article's topic (*NewsHydrogen, -Ammonia, -Methanol, -Efuel, -EV*)
- Five binary variables for the company's technological classification (*CompanyHydrogen, -Ammonia, -Methanol, -Efuel, -EV*)

These binary variables now serve to create five binary *Match*-variables, which equal 1 when the news article's technology topic matches the company's technology (*MatchHydrogen, -Ammonia, -Methanol, -Efuel, -EV*). For instance, *MatchHydrogen* is 1 when the binary variables *NewsHydrogen* and *CompanyHydrogen* are both 1, otherwise *MatchHydrogen* is 0.

PART 2.2: Panel Regression Analysis – Matching Technology

This section uses the same regression model as in PART 1.2, but analyzes each clean technology subgroup individually. It divides the total dataset into five subsets, with each subset representing a match between the news topic and the company's technology focus. For example, for the hydrogen technology subgroup, only observations where the binary variable *MatchHydrogen* equals 1 are included in the analysis. This is then also done for the other four technology subgroups.

This panel regression model corresponds to the exact same model as in PART 1.2, but is repeated for each of the five technologies.

$$CAR\_0\_10_{i,n} = \sum_{k=1}^{2} \beta_{1k}\, Tone_{k,n} + \sum_{l=1}^{3} \beta_{2l}\, Politics_{l,n} + \sum_{k=1}^{2}\sum_{l=1}^{3} \beta_{3kl}\, (Tone_{k,n} \times Politics_{l,n}) + \gamma_{i,n} + \varepsilon_{i,n}$$

5. Analysis

5.1 Exploratory Data Analysis

5.1.1 CompanyData

Out of the total sample of 108 companies in the CompanyData dataset, Figure 1 shows that 91 companies are based in the U.S., while 17 companies are based in the UK. Figure 2 represents the industry classification of the companies, which reveals a concentration in Energy Technologies (34 companies), Energy Storage (20 companies), and Electric Vehicles (17 companies).

Figure 1: Distribution of U.S. & UK Companies

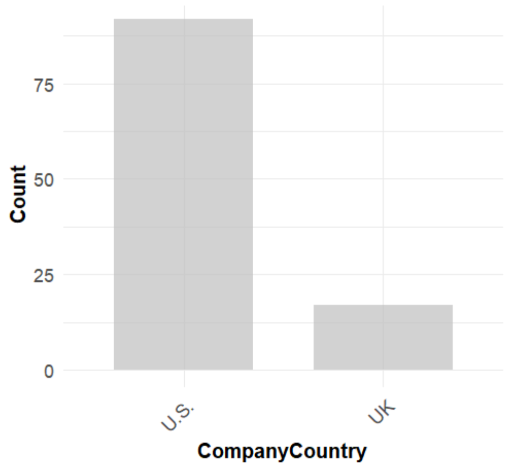
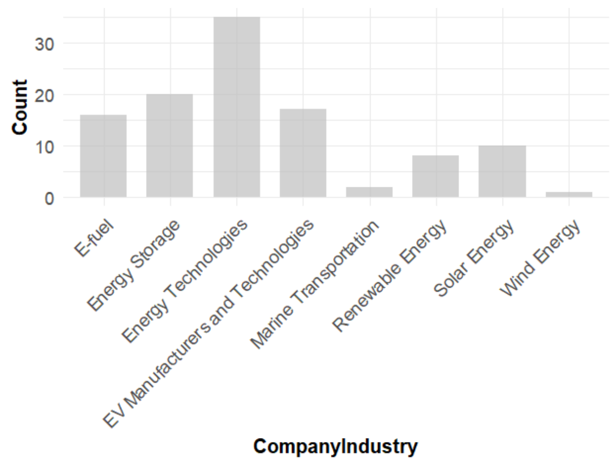


Figure 2: Distribution of Company Industries



Source: Verdonck, 2025e; using data from S&P Global Market Intelligence, n.d.

5.1.2 NewsData

The news article dataset comprises 1558 clean technology-related headlines collected between 2007 and 2022.

The distribution of political alignment in the news sample (Figure 3 and 4) shows a dominance of center-aligned sources, primarily driven by CNBC (722 news articles), with fewer articles from far-right and right-leaning outlets. FOX News, being the only far right newspaper, has only 64 articles to its name, while the Daily Mail and the New York Post, both of right inclination, have 111 news articles to their count together. The other five newspapers of left inclination provide 522 news articles to the dataset. There is a significant disequilibrium between the different political alignments, especially between center and (far) right. This could partially be explained by the fact that left-leaning and centrist outlets generally demonstrate greater receptivity to publish about clean technology topics than right-wing media.

Sentiment analysis identifies a balanced distribution between neutral (659) and positive (565) sentiment headlines, while negative (334) sentiment articles account for approximately 21% of the dataset (Table 2). Table 2 shows that clean technology news articles from Left- and Center-oriented newspapers are primarily neutral (41% and 45%, respectively), closely followed by positive (38%

and 39%, respectively), while Right- and Far Right-oriented newspapers publish primarily negative (53% and 44%, respectively) of negatively toned news articles.

Figure 3: Distribution of Newspapers

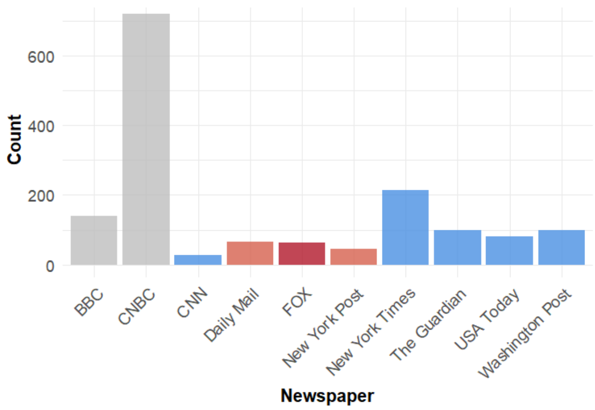
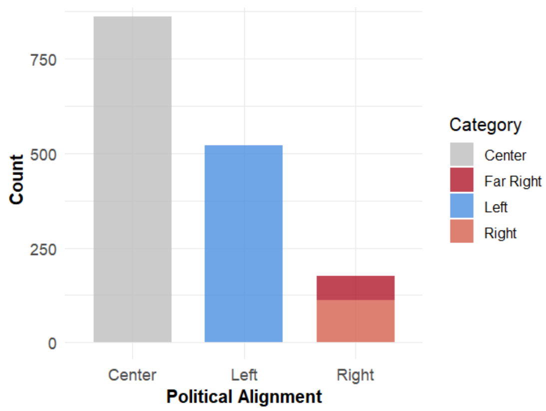


Figure 4:Distrubution of Political Alignment



Source: Verdonck, 2025a; using data from Statista, 2025c; Ground News, n.d.; Kaggle, 2023

Table 2: Distribution of Political Alignment and Sentiments

Sentiment/Politics	Left	Center	Right	Far Right	Total
Negative	110 (0.21)	137 (0.16)	59 (0.53)	28 (0.44)	334 (0.21)
Neutral	215 (0.41)	392 (0.45)	30 (0.27)	22 (0.34)	659 (0.42)
Positive	196 (0.38)	333 (0.39)	22 (0.20)	14 (0.22)	565 (0.36)
Total	521 (0.33)	862 (0.55)	111 (0.07)	64 (0.04)	1558 (1.00)

Source: Verdonck, 2025a; using data from Statista, 2025c; Ground News, n.d.; Kaggle, 2023

5.2 Hypothesis Testing

PART 1: Influence of Media Coverage on Stock Returns

PART 1.1: Event Study – Significance of the Mean CAR

In order to examine whether a given news announcement about clean technology causes abnormal stock returns for the companies in question, a CAR analysis is conducted for every news article publication date per company. First, the market model and CAPM parameters (i.e., α and β) are estimated using a 200-day pre-event estimation window [-220, -21], thereby capturing normal return behavior for each company prior to the news publication. Next, daily abnormal returns are computed over a large window of [−20, 60] trading days around every event date, enabling a subsequent partition of these abnormal returns into shorter subwindows: [-5, 0], [-5, 10], [0,10] and [11,20]. The CARs are then obtained by summing the daily ARs within each subwindow for every firm after every event date, leading to a dataset of 128372 rows, containing panel data of each CAR for each company and subwindow per event date. The CAR estimates obtained from both the market model and the CAPM appear to be nearly identical across event windows. Given its theoretical advantages, the CAPM-based CARs are selected for all subsequent analyses.

Summary Statistics

Since especially the immediate effect of a headline publication on the stock market return is of interest, the CAR over the subwindow [0,10], *CAR_0_10*, is focused on.

In Table 3, the summary statistics of the mean CAR over the subwindow [0,10] show that the data is highly skewed due to a positive outlier, with the maximum value being 65.66 (or 6566%). This extreme value distorts the distribution, inflating the mean (0.28%) and the standard deviation (43.2%), which suggests substantial variability in CARs across companies. The median is slightly negative (-0.14%), indicating that the majority of the companies experienced small negative abnormal returns, while the positive skew indicates the one firm, Blink Charging Co, that has very large positive returns at the start of going public. While this outlier is not due to erroneous data and reflects real stock price changes for the company, it could disproportionately influence the mean CAR, subsequent t-tests and further analysis.

This motivates winsorizing that CAR data over the subwindow [0,10] in order to reduce the impact of the outlier. Winsorizing is considered a robust statistical method that adjusts for the influence of extreme observations while retaining the data in the analysis. In this case, values above the 99th percentile were capped to the value at the 99th percentile (Wilcox, 2022). This transformation ensures that the majority of the data remains intact, while the most extreme values are brought within a more reasonable range. After winsorizing the data, the maximum CAR value is 3.99 (or 399%), reducing the impact of the outlier. The mean has decreased to 0.0366%, and the standard deviation dropped to 13.8%, indicating less variability and a more stable dataset.

Table 3: Summary Statistics for the initial and the winsorized CARs [0,10]

	Min	1st Qu.	Median	Mean	3rd Qu.	Max	sd
CAR_0_10	-1.67366	-0.04912	-0.00138	0.00280	0.03832	65.65901	0.4319728
CAR_0_10_winsor	-1.67366	-0.04912	-0.00138	0.00037	0.03832	3.99731	0.1379177

Source: Verdonck, 2025a-e; using data from S&P Global Market Intelligence, n.d.; Bloomberg L.P., n.d.; Federal Reserve Bank of St. Louis., n.d.

In the following analysis, the winsorized CARs over the subwindow [0,10] will be considered to ensure prudency.

Event Study Results

To evaluate whether the CARs across companies deviate significantly from zero and to examine their directional tendencies, a binomial test and a series of t-tests are conducted.

A binomial test (Table 4) was conducted to determine whether the majority of companies experienced positive or negative CARs within the [0,10] event window. This test evaluates the proportion of companies with mean CARs greater than zero and assesses whether this proportion significantly differs from 50%. Although 57% of observations exhibited positive CARs, the binomial test failed to reject the null hypothesis that this proportion equals 50%.

Table 4: Directional Significance of CAR_0_10

Window	n_total	n_positive	proportion_positive	Significance
[0, 10]	108	62	0.57	

Source: Verdonck, 2025a-e; using data from S&P Global Market Intelligence, n.d.; Bloomberg L.P., n.d.; Federal Reserve Bank of St. Louis., n.d.

After conducting t-tests on the initial and winsorized mean CAR, the initial mean CAR for the subwindow [0,10] is significant and positive (Table 5), which reflects a general positive abnormal return in that period. However, this significance diminishes when analyzing the winsorized mean CAR (Table 6). This result is not unexpected, since the winsorization process, by capping extreme positive CAR values, reduces the influence of large, outlier values that could disproportionately affect the mean. In doing so, it leads to a less pronounced positive mean and a loss of statistical significance.

Table 5: Significance initial mean CAR

	Window	Mean_CAR	Significance
1	[-5, 10]	0.00208	*
2	[-5, 0]	-0.00034	
3	[0, 10]	0.00279	**
4	[11, 20]	0.00118	*

Table 6: Significance winsorized mean CAR

	Window	Mean_CAR_winsor	Significance
1	[-5, 10]	-0.00037	
2	[-5, 0]	-0.00085	***
3	[0, 10]	0.00037	
4	[11, 20]	0.00068	*

Source: Verdonck, 2025a-e; using data from S&P Global Market Intelligence, n.d.; Bloomberg L.P., n.d.; Federal Reserve Bank of St. Louis., n.d.

It is important to note that this is an aggregated result, and the CARs in this window range from both positive to negative values, represent companies across all clean technologies and news articles with different content characteristics. This aggregated measure may not fully capture the complexity of individual market reactions. Therefore, the loss of significance in the winsorized version of the CAR is not necessarily problematic. The following regression analyses investigate the moderating role of media tone, political orientation, and technological alignment. In those disaggregated models, the CAR may still be significant within certain subgroups, which provides a more granular understanding of the observed effects.

PART 1.2: Panel regression Analysis – Influence of News Sentiment and Politics

Summary Statistics of Variables Included in Panel Regression

<u>Sentiment, Politics, Sentiment x Politics</u>

Table 7 shows that the panel data has a fairly balanced distribution of neutral and positive sentiments (42% and 37%, respectively). The negative sentiment is the least frequent. Center-leaning newspaper are more prevalent (61%), with a smaller proportion of left- and right-oriented newspapers (29% and 10%, respectively). Regarding the interaction between sentiment and politics, Left- and Center-oriented newspapers publish mostly neutral news articles, while right- and far right-oriented newspapers publish mostly news articles with negative sentiment. These results of the panel dataset align with the previous Table 2, indicating that the panel dataset represents that original newspaper dataset very well.

Table 7: Distribution of Sentiment and Politics

Sentiment/Politics	Left	Center	Right	Far Right	Total
Negative	7788 (0.21)	12320 (0.16)	4434 (0.51)	1939 (0.45)	26481 (0.21)
Neutral	14879 (0.41)	35654 (0.45)	2469 (0.28)	1302 (0.30)	54304 (0.42)
Positive	14014 (0.38)	30704 (0.39)	1800 (0.21)	1069 (0.25)	47587 (0.37)
Total	36681 (0.29)	78678 (0.61)	8703 (0.07)	4310 (0.03)	128372 (1.00)

Source: Verdonck, 2025a-e; using data from Statista, 2025c; Ground News, n.d.; Kaggle, 2023; S&P Global Market Intelligence, n.d.; Bloomberg L.P., n.d.; Federal Reserve Bank of St. Louis., n.d.

<u>Control variables</u>

- News control variable:

NewsCountry: There is a majority of observations of news articles being published by U.S. newspapers (105209 or 82%) compared to UK newspapers (23163 or 18%)

- Company control variables:

USCompany: There is a clear majority of observations concerning U.S. companies (107200 or 84%) compared to UK companies (21171 or 16%) in the dataset.

CompanyIndustry: Companies within Energy Technologies represent the largest share (34%) compared the other industries, E-fuel (18%), Energy Storage (14%) and EV Manufacturers and Technologies (15%) represent about equal shares in the dataset (Table 8).

Table 8: CompanyIndustry

CompanyIndustry	Amount and Proportion
E-fuel	23482 (0.18)
Energy Storage	18302 (0.14)
Energy Technologies	44914 (0.35)
EV Manufacturers and Technologies	19041 (0.15)
Marine Transportation	2793 (0.02)
Renewable Energy	9059 (0.07)
Solar Energy	9664 (0.08)
Wind Energy	1117 (0.01)
Total	128372 (1.00)

Source: Verdonck, 2025e; using data from S&P Global Market Intelligence, n.d.

Total_Assets_log and *Total_Employees_log*: As the mean of both *Total_Assets* and *Total_Employees* is much bigger than the median and the maxima show extreme right skewness, the log is taken of both variables (Table 9).

Table 9: Total_Assets(_log) and Total_Employees(_log)

Variable/statistics	Min	1st Qu.	Median	Mean	3rd Qu.	Max
Total_Assets	0	78.10	440.40	17790	4877.80	641343
Total_Employees	1	121	990	10152	8450	225000
Total_Assets_log	0	4.37	6.09	6.38	8.49	13.37
Total_Employees_log	0	4.80	6.90	6.83	9.04	12.32

Source: Verdonck, 2025a-g; using data from Bloomberg L.P., n.d.

- Time control variable

Year: Most clean technology articles have been published in the year 2022 (38%) (Table 10).

Table 10: Year

Year	Amount and Proportion
2022	49114 (0.38)
2021	30736 (0.24)
2020	8170 (0.06)
2017	6964 (0.05)
2016	5190 (0.04)
2019	4894 (0.04)

Source: Verdonck, 2025a-e; using data from Bloomberg L.P., n.d.

Panel Regression Results

As a reminder of the three hypotheses tested in this section,

H2 – Tone Effect: The tone of newspaper articles on clean technologies influences the stock market returns of clean tech companies.

H3 – Political Alignment Effect: The political alignment of the publishing newspaper influences the stock market returns of clean tech companies.

H4 – Interaction Effect: The interaction between the tone of the article and political alignment of the publishing newspaper influences the stock market returns of clean tech companies.

This comprehensive fixed effects panel regression model tests the explanatory power of media characteristics on CARs.

$$CAR\_0\_10_{i,n} = \sum_{k=1}^{2} \beta_{1k}\, Tone_{k,n} + \sum_{l=1}^{3} \beta_{2l}\, Politics_{l,n} + \sum_{k=1}^{2}\sum_{l=1}^{3} \beta_{3kl}\, (Tone_{k,n} \times Politics_{l,n}) + \gamma_{i,n} + \varepsilon_{i,n}$$

Table 11 shows the results of three panel regressions, regressing sentiment, political alignment, and their interaction on the dependent variable, the winsorized 10-day CARs of clean technology companies[1]. The first regression model (1) includes only the news sentiment (*Tone*) variable, the second regression model (2) includes only the political alignment (*Politics*) variable, the third regression model (3) is represents the comprehensive model, which includes the *Tone*, the *Politics* and the interaction variable between *Tone* and *Politics*. All models include a comprehensive set of control variables related to the news (*NewsCountry*), company (*CompanyCountry, CompanyIndustry, Total_Assets_log, Total_Employees_log*) and time (*Year*). All independent variables in the regression have been tested for multicollinearity by using the Variance Inflation Factor (VIF).

[1] These models were repeated with the initial CAR_0_10 as the dependent variable (not winsorized), which led to almost identical results and the same interpretation.

Table 11: Panel Regression Results for Media Content Effects[2]

	Dependent variable:		
	CAR_0_10_winsor		
	Tone Effect	Politics Effect	Individual and Interaction Effect
	(1)	(2)	(3)
SentimentNegative	0.004***		0.008***
	(0.001)		(0.001)
SentimentPositive	0.002**		0.004***
	(0.001)		(0.001)
PoliticsFar Right		0.005**	−0.003
		(0.002)	(0.004)
PoliticsLeft		0.003***	0.006***
		(0.001)	(0.001)
PoliticsRight		−0.002	0.009***
		(0.002)	(0.003)
SentimentNegative:PoliticsFar Right			0.005
			(0.005)
SentimentPositive:PoliticsFar Right			0.017***
			(0.006)
SentimentNegative:PoliticsLeft			−0.006**
			(0.002)
SentimentPositive:PoliticsLeft			−0.006***
			(0.002)
SentimentNegative:PoliticsRight			−0.017***
			(0.004)
SentimentPositive:PoliticsRight			−0.020***
			(0.004)
News Control Variables	Included	Included	Included
Company Control Variables	Included	Included	Included
Time Control Variables	Included	Included	Included
Observations	112,306	112,306	112,306
R^2	0.011	0.011	0.011
Adjusted R^2	0.010	0.010	0.010
Note:			*p<0.1; **p<0.05; ***p<0.01

Source: Verdonck, 2025a-g; using data from Statista, 2025c; Ground News, n.d.; Kaggle, 2023; S&P Global Market Intelligence, n.d.; Bloomberg L.P., n.d.; Federal Reserve Bank of St. Louis., n.d.

In Model (1), both sentiment variables are statistically significant: negative sentiment is associated with a CAR increase of 0.4 percentage points, while positive sentiment yields a smaller but still significant increase of 0.2 percentage points, relative to the baseline of neutral sentiment. This implies that sentiment, both positive and negative, can drive short-term investor reactions.

Model (2) includes only political orientation. Compared to the baseline of politically centrist outlets, articles from far-right and left-leaning sources are associated with significantly higher CAR increases of 0.5 percentage points and 0.3 percentage points, respectively. Articles from right-leaning outlets, however, do not differ significantly from the baseline. This indicates investors react more positively to news from political extremes, especially far-right sources.

[2] Because the fixed effects panel regression requires a unique (date × company) index, any dates on which a company has multiple headlines are automatically dropped. As a result, the regression dataset contains fewer observations than the total number (128372)

Model (3) presents the combined effects of sentiment and political alignment on CARs relative to the baseline of neutral sentiment and centrist politics. Negative sentiment in far-right outlets is associated with an increase of approximately 1.0 (0.5 + 0.8 – 0.3) percentage point in CARs, while positive sentiment in far-right outlets corresponds to an increase of about 1.8 (1.7 + 0.4 – 0.3) percentage points. For left-leaning outlets, negative and positive sentiments yield increases of roughly 0.8 (– 0.6 + 0.8 + 0.6) and 0.4 (– 0.6 +0.4 + 0.6) percentage points, respectively. In contrast, negative sentiment in right-leaning outlets results in a negligible net effect close to zero (– 1.7 + 0.8 + 0.9) , whereas positive sentiment in right-leaning outlets is associated with a decline of approximately 0.7 (–2.0 + 0.4 + 0.9) percentage points in CARs.

The adjusted R^2 across all models remains low (0.01), which is expected given the presence of other influential market drivers. Nevertheless, the models are robustly specified and the statistically significant coefficients indicate that sentiment, political bias, and their interaction affect short-term stock market reactions to clean technology news. These findings suggest that the investor response to media sentiment varies depending on the political orientation of the news source, with far-right and left-leaning outlets generally amplifying positive returns, while right-leaning outlets attenuate or reverse the effect of positive sentiment. These findings align with media slant literature (Gentzkow & Shapiro, 2010) and reinforce the theoretical importance of limited attention and sentiment-driven behavior (Kahneman, 1973; Barber & Odean, 2008).

Now that the link has been established between news coverage, characterized by sentiment and political inclination, and abnormal stock returns of clean technology companies, the next part of the analysis analyzes technology-specific sensitivities for five specific technologies.

PART 2: Subgroup Comparison – Technology-Specific Sensitivities

The third and final part of the analysis explores heterogeneous effects across technological subgroups. This section tests whether technological alignment between news topics and firm technologies amplify stock return responses.

As a reminder of Hypothesis 5 tested in this section,

Hypothesis 5 (Matching Technology Effect): *News articles covering a specific technology have a stronger effect on companies active in that same technology.*

PART 2.1: Event Study – Significance of the Mean CAR by Matching Technology[3]

Summary Statistics

Table 12: Amount and Proportions of News Articles and Companies present in the Ammonia, Methanol, Hydrogen, E-fuel and/or EV sector

	Hydrogen	Ammonia	Methanol	Efuel	EV
News	18047 (14.1%)	883 (0.69%)	1319 (1.03%)	5200 (4.12%)	15506 (12.08%)
Company	31575 (24.6%)	17643 (13.7%)	16940 (13.2%)	26698 (20.8%)	19041 (14.8%)
MatchNewsCompany	4520 (3.52%)	157 (0.12%)	192 (0.15%)	1269 (0.99%)	2354 (1.83%)

Source: Verdonck, 2025a-g; using data from Statista, 2025c; Ground News, n.d.; Kaggle, 2023; S&P Global Market Intelligence, n.d.; Bloomberg L.P., n.d.; Federal Reserve Bank of St. Louis., n.d.

[3] These t-tests were repeated with the initial mean CAR_0_10 (not winsorized), which led to almost identical results and the same interpretation.

Table 12 reports the proportion of news articles and companies that are associated with each of the five clean technology sectors considered in this part of the study. The proportions are computed relative to the total dataset of 128372 lines. Note that the values do not sum to 100%, as there are observations in the dataset that do net enter into any of the five technology subgroups. The most frequently mentioned technology in the media is hydrogen, appearing in 14.1% of the observations, followed by electric vehicles at 12.08%. In contrast, ammonia and methanol receive little media attention, with proportions of only 0.69% and 1.03%, respectively. On the company side, the distribution is more balanced, though still led by hydrogen (24.6%), followed by E-fuel (20.8%) and EV (14.8%). Ammonia and methanol account for a substantially larger share of companies (13.7% and 13.2%) than of news articles, suggesting a potential underrepresentation of these technologies in media coverage relative to their presence in the market.

Event Study Results

The following analysis examines the results of t-tests conducted to assess the effects of specific technology themed news articles on the CAR of companies within the different clean technology sectors of Hydrogen, Ammonia, Methanol, E-fuel and Electric Vehicles. The results allow to investigate sector-specific sensitivities to news coverage, the possible existence of a substitution effect between technologies, and whether the alignment between the news topic and the company's core business plays a role in influencing CAR.

Table 13: Mean CAR_0_10_winsor and Significance depending on the News- and Company Technology

NewsType	CompanyHydrogen	CompanyAmmonia	CompanyMethanol	CompanyEfuel	CompanyEV
NewsHydrogen	0.0072 ***	0.0087 ***	0.0046 **	0.0073 ***	-0.001
NewsNoHydrogen	0.0019 **	0.0015 *	0	0.0018 **	0.0015
NewsAmmonia	0.0015	0.0033	0.0127	0.0026	0.0111
NewsNoAmmonia	0.0026 ***	0.0025 ***	0.0006	0.0026 ***	0.0010
NewsMethanol	0.0250 ***	0.0239 ***	0.0213 ***	0.0243 **	0.0381 **
NewsNoMethanol	0.0024 ***	0.0023 ***	0.0004	0.0024 ***	0.0007
NewsEfuel	0.0061 **	0.0056	0.0045	0.0056	0.0284 ***
NewsNoEfuel	0.0025 ***	0.0023 ***	0.0005	0.0025 ***	0
NewsEV	-0.0002	-0.002	-0.0017	-0.0018	-0.0033
NewsNoEV	0.003 ***	0.0031 ***	0.001	0.0032 ***	0.0017

Source: Verdonck, 2025a-g; using data from Statista, 2025c; Ground News, n.d.; Kaggle, 2023; S&P Global Market Intelligence, n.d.; Bloomberg L.P., n.d.; Federal Reserve Bank of St. Louis., n.d.

Table 13 presents the winsorized mean CARs over the [0,10] subwindow for clean technology companies, segmented by both the technology focus of the news article (rows) and the company's technology classification (columns). Each cell reports the mean CAR followed by the significance stars from a t-test evaluating whether the mean CAR differs significantly from zero. 'NewsX' rows represent news headlines directly related to a given technology, while 'NewsNoX' rows represent all news headlines not related to that technology. Dark grey cells denote instances where the company's technology aligns with the topic of the news article. Light grey cells indicate instances where the company's technology does not align at all with the topic of the news article.

Methanol and Hydrogen are the only two technologies where the alignment between news- and company topic has a significant and positive mean CAR, 2.13% and 0.72% respectively. Interestingly, non-matching technology news also significantly influences CARs, as portrayed in the light grey cells for Ammonia, Hydrogen and E-fuel, with mean CARs of 0.25%, 0.19%, and 0.25% respectively.

When looking at interactions between different technologies, Methanol news appears to benefit ammonia (2.39%), hydrogen (2.5%), e-fuel (2.43%), and EV (3.81%) companies significantly, suggesting spillover effects and perceived technological complementarity. Similar but weaker patterns emerge for hydrogen and e-fuel news. In contrast, EV-related news has no significant effect on companies across different technologies. Zooming in closer on the potential substitution of complementary effect between methanol and ammonia (World Bank, 2024; IRENA, 2021), the results indicate that methanol news significantly benefits both methanol and ammonia companies positively, while ammonia news does not significantly influence methanol or ammonia companies.

Overall, this table highlights that matched news coverage tends to cause stronger stock market reactions than unrelated news, but the strength and direction of the effect vary across technologies. It also reveals the presence of cross-technology spillover effects, particularly from methanol news, and potential saturation effects in the EV sector. However, investor expectations, other external factors or news specific factors, such as the news outlet's political orientation or the sentiment of the news article may moderate the impact of the news on the different subgroups of companies. To capture this additional layer of complexity, the following section extends the analysis by estimating a series of panel regressions for each clean technology subgroup separately. These models include interaction terms for sentiment and political alignment, allowing for a more nuanced understanding of how the content and ideological framing of news articles moderate the effect of technology-specific news on firm-level CARs.

PART 2.2: Panel Regression Analysis – Matching Technology

For each clean technology – Hydrogen, Ammonia, Methanol, E-fuel, and Electric Vehicles – a separate panel regression model is estimated. The models incorporate interaction terms between media sentiment and political orientation, with the dependent variable being the winsorized CAR of each subgroup over the [0,10] subwindow. This approach allows for a more granular analysis of the conditions under which news coverage generates statistically and economically meaningful investor responses within each technology category. The regression uses a fixed effects specification ("within" model) to control for unobserved, time-invariant company-specific heterogeneity, with 'individual' effects specified to account for company-specific differences in their response to news.

Recall the panel regression model, estimated for each technology:

$$CAR\_0\_10_{i,n} = \sum_{k=1}^{2} \beta_{1k}\, Tone_{k,n} + \sum_{l=1}^{3} \beta_{2l}\, Politics_{l,n} + \sum_{k=1}^{2}\sum_{l=1}^{3} \beta_{3kl}\, (Tone_{k,n} \times Politics_{l,n}) + \gamma_{i,n} + \varepsilon_{i,n}$$

All independent variables in the regression models have been tested for multicollinearity by using the Variance Inflation Factor (VIF). In order to avoid multicollinearity, the five different models per technology match do not all contain all control variables.

- News control variable: *NewsCountry* (Model 1, 2, 3, 4, 5)

- Company control variables: *CompanyCountry* (Model 1, 2, 3, 4, 5), *CompanyIndustry* (Model 1, 2, 3, 4), *Total_Assets_log* (Model 1, 2, 3, 4, 5), *Total_Employees_log* (Model 1, 2, 3, 4, 5)
- Time control variables: *Year* (Model 1, 5)

Table 14: CAR_0_10_winsor Panel Regression for Every Technology

	Dependent variable:				
	CAR_0_10_winsor				
	Hydrogen	Ammonia	Methanol	E-fuel	EV
	(1)	(2)	(3)	(4)	(5)
SentimentNegative	−0.017**	0.025	0.047	0.013	0.036***
	(0.008)	(0.031)	(0.038)	(0.009)	(0.012)
SentimentPositive	0.003		0.054*	−0.004	0.031***
	(0.006)		(0.030)	(0.010)	(0.010)
PoliticsFar Right	0.044			0.002	0.008
	(0.033)			(0.020)	(0.034)
PoliticsLeft	0.023***	−0.013		−0.005	0.029*
	(0.008)	(0.028)		(0.011)	(0.015)
PoliticsRight	−0.001	−0.012	0.054	0.005	0.068***
	(0.015)	(0.033)	(0.035)	(0.012)	(0.021)
SentimentNegative:PoliticsFar Right	−0.040				0.078
	(0.035)				(0.061)
SentimentPositive:PoliticsFar Right	−0.053				−0.023
	(0.039)				(0.048)
SentimentNegative:PoliticsLeft	−0.011				−0.017
	(0.012)				(0.034)
SentimentPositive:PoliticsLeft	−0.036***				−0.045**
	(0.011)				(0.022)
SentimentNegative:PoliticsRight	−0.008				−0.129***
	(0.019)				(0.034)
SentimentPositive:PoliticsRight	0.065**				−0.106***
	(0.030)				(0.031)
News Control Variables	Included	Included	Included	Included	Included
Company Control Variables	Included	Included	Included	Included	Included
Time Control Variables	Included	Not Included	Not Included	Not Included	Included
Observations	4,031	144	175	1,123	2,183
R^2	0.048	0.076	0.066	0.014	0.038
Adjusted R^2	0.037	−0.049	−0.028	−0.008	0.021
Note:					*p<0.1; **p<0.05; ***p<0.01

Source: Verdonck, 2025a-g; using data from Statista, 2025c; Ground News, n.d.; Kaggle, 2023; S&P Global Market Intelligence, n.d.; Bloomberg L.P., n.d.; Federal Reserve Bank of St. Louis., n.d.

Table 14[4] presents the interaction results of five fixed effects panel regressions examining how sentiment and political orientation influence the *CAR_0_10_winsor* for each clean technology. Only the models for Hydrogen and EV companies indicate explanatory power. Their adjusted R^2 are small but significant (0.037 and 0.021 respectively), while the adjusted R^2 for Ammonia, Methanol and E-fuels are negative, indicating a very bad fit. This bad fit could not be improved upon by removing or adding control variables. It can be due to the very limited matched data between news article technology topic and company technology (see Table 12). Nonetheless, for methanol companies,

[4] For the technologies Ammonia, Methanol and E-fuels, the panel regression model does not include the interaction variables (due to multicollinearity issues). For Hydrogen and EV, the model only includes the interaction variables, since the individual effects of Sentiment and Politics did not add substantially change any interpretation.

positive sentiment shows a significant positive effect (5.4 percentage points), aligning with earlier evidence that methanol companies benefit from favorable media exposure.

For hydrogen companies (Column 1), negative sentiment significantly decreases CAR by 1.7 percentage points, compared to the baseline neutral sentiment, while left-leaning outlets are associated with a positive effect (2.3 percentage points) compared to centrist politics. However, when left-leaning political alignment is paired with positive sentiment, the effect decreases by 1 percentage point (0.03 + 2.3 – 3.6) compared with neutral sentiment and centrist politics, suggesting that investors may discount favorable tone from left-leaning sources in this sector. The interaction effect between positive sentiment and right politics leads to a high positive effect of 6.9 percentage points (6.5 + 0.3 – 0.1), suggesting that when clean technology news is framed positively by right-leaning outlets, investors interpret it as a particularly strong signal, perhaps because it represents a contrarian opinion that enhances the credibility of the underlying innovation.

For Electric Vehicle companies (Column 5), sentiment effects are particularly strong. Negative sentiment increases CARs by 3.6 percentage points, and positive sentiment also has a strong positive effect (3.1 percentage points) compared to the baseline Neutral Sentiment. Compared to centrist politics, the left increases CARs by 2.9 percentage points and right increases CARs by 6.8 percentage points. The interaction terms reveal heterogeneity in how investors react to sentiment depending on political orientation, compared to the neutral sentiment and centrist politics baseline. Positive sentiment from left-leaning outlets yields a CAR increase of 1.5 percentage points (–4.5 + 3.1 + 2.9), whereas positive sentiment from right-leaning outlets results in a decrease of 0.7 percentage points (–10.6 + 3.1 + 6.8). Similarly, negative sentiment in right-leaning outlets produces a drop of 2.5 percentage points (–12.9 + 3.6 + 6.8). These patterns indicate that, although both positive and negative news generally drive short-term returns, sentiment interactions with left-leaning outlets tend to produce more positive CARs than those with right-leaning outlets.

5.3 Analysis Conclusion

PART 1 analyzes the broad influence of media coverage on clean technology stock returns. The event study results in PART 1.1 do not provide unambiguous support for Hypothesis 1 (H1) – the broad winsorized 10-day mean CAR is not statistically significant – but rejection would be too precipitate. Legal newspaper coverage, despite its capacity to simplify complex innovations and signal legitimacy (Rogers, 2003), causes heterogeneous investor responses depending on company technologies and news context. This is analyzed in subsequent analysis.

The fixed effects panel regression in PART 1.2 validates Hypotheses 2–4, demonstrating that both sentiment (H2), political orientation (H3) and its interaction (H4) shape returns. Both negative and positive clean technology headlines independently boost CARs by 0.4 and 0.2 percentage points versus neutral coverage, demonstrating that any sentiment attracts investor attention. Articles from far-right outlets raise CARs by 0.5 percentage points and left-leaning outlets by 0.3 percentage points compared to centrist politics. Finally, when interacting sentiment with political orientation, far-right coverage amplifies both negative (1.0 percentage points) and positive (1.8 percentage points) effects, left-leaning coverage causes CARs increase for negative (0.8 percentage points) and positive (0.4 percentage points) tones, and right-leaning coverage neutralizes negative tone and even reverses positive tone into a − 0.7 point drop. Despite statistical significance, the modest R^2 (0.011) suggests that media framing explains only a small portion of CAR variation, emphasizing the role of broader market forces.

The technology-specific analysis in PART 2 confirms Hypothesis 5 (H5): media effects vary by technology type and news-company technology alignment. The technology-match event study in PART 2.1 reveals that only methanol and hydrogen news generate significantly positive abnormal returns for firms operating in the corresponding technologies (2.13% and 0.72%, respectively), while non-matched coverage also yields modest yet significant effects for ammonia, hydrogen, and e-fuel firms (0.25%, 0.19%, and 0.25%). Methanol-focused articles exhibit pronounced spillover benefits across multiple sectors – ammonia (2.39%), hydrogen (2.50%), e-fuel (2.43%), and EV (3.81%) – whereas EV CARs across sector do not appear to be significantly different from zero.

The technology-specific panel regressions in PART 2.2 reveal that only the Hydrogen and EV models achieve sufficient statistical robustness (adjusted R^2 of 0.037 and 0.021, respectively), whereas Ammonia, Methanol, and E-fuel models suffer from inadequate fit due to limited matched observations. In the hydrogen sector, negative sentiment in centrist outlets decreases CARs by 1.7 percentage points, and left-leaning coverage alone adds 2.3 percentage points, but this increase is offset when positive tone interacts with left-wing politics. Notably, positive framing by right-leaning outlets yields a pronounced 6.9 percentage points lift, suggesting contrarian credibility. For EV firms, both negative (3.6 percentage points) and positive (3.1 percentage points) sentiment markedly boost returns, yet their interaction with outlet ideology yields divergent effects: left-leaning interactions yield modest positive CARs (1.5 percentage points) while right-leaning interactions reverse positive tone into a –0.7 percentage points decline and exacerbate negative tone into a –2.5 percentage points decrease.

These results indicate that the same sentiment-politics dynamics occur in PART 1 (the broad dataset) and PART 2 (the dataset divided by technology subgroup). Media visibility, sentiment framing, and ideological context jointly shape short-term clean technology investment dynamics, but their ultimate impact depends critically on the specific technology focus and news-company alignment.

6. Conclusion, Limitations and Further Research

6.1 Conclusion

As a reminder of the central research question,

‘*How does investor attention to clean technology news coverage, conditioned by media tone and political orientation, influence the investment into clean technology alternatives?*’

Legal media coverage acts not only as a reflection of clean technology trends but as a mean in shaping them. Building on Rogers’s (2003) diffusion framework and Kemp & Volpi’s (2008) insights on clean technology adoption, strategic legal media communication can help accelerate capital allocation toward clean technologies, thereby supporting the diffusion of innovations essential for decarbonization. The technology subgroup analysis (PART 2) contrasts electric vehicles – a more mature technology – and hydrogen, ammonia, methanol and e-fuels –emerging technologies that face higher uncertainty (Comin & Nanda, 2019; Lanteri & Rampini, 2025). Consistent with the ‘two-speed’ transition (International Energy Agency, 2025), the event study (PART 2.1) shows that mean CARs are significant for emerging technologies but less so for electric vehicles. This indicates that legal media coverage can be more impactful in informing investors about emerging technologies than about more mature technologies. However, the results also caution against overreliance on media visibility alone: the small explanatory power of the panel regression models indicate that while media coverage is impactful, it remains a complement rather than a substitute to sound fundamentals, supportive policy, and scalable business models.

Investor attention to clean-technology news is driven as much by media tone as by political orientation: both positive and negative headlines significantly boost short-term abnormal returns versus neutral coverage, confirming that any sentiment attracts investors. Unexpected tone-political orientation combinations (e.g. positive sentiment in right-leaning media) yield the largest return outcomes.

From a policy perspective, these findings suggest that governments and advocacy groups must recognize the media's dual role: as a signal amplifier for market confidence, and as a potential hurdle that can hinder adoption through limited coverage, as seen for ammonia, methanol and e-fuels more broadly. For investors and companies, the research provides guidance on how strategic media engagement might enhance capital access. For academics, it opens new avenues to explore the intersection of information dynamics and sustainable finance.

6.2 Limitations and Opportunities for Further Research

While this thesis advances understanding of how media coverage, its sentiment and political orientation influence stock market reactions and clean technology adoption, several limitations give rise to opportunities for further research.

A first limitation is that the newspaper dataset only extends through 2022, due to data limitation constraints. Given that global clean technology investments increased from $1.5 trillion in 2022 to over $2 trillion in 2024 (Statista, 2025a), incorporating more recent data could provide insights into the latest market dynamics and trends.

Secondly, the study focuses on traditional legal newspapers as media sources, excluding social media, financial analyst reports, or other digital channels that increasingly shape investor attention, especially among retail investors. Different investor types may respond differently to media types, and the omission of these channels restricts a comprehensive understanding of the media-investor relation in clean technology finance.

Third, this study concentrates only on U.S.- and UK-listed clean technology companies and newspapers from these countries. This geographical focus limits the generalizability of the findings to other regions where media, investor behavior, and clean technology adoption dynamics may differ significantly. Therefore, extending this research to emerging markets and Europe may be beneficial for a comparative understanding on how news coverage impacts clean technology investments globally.

Fourth, the analysis examines short-term CARs within a 10-day window following news publication. While appropriate for capturing immediate investor reactions, this approach does not address longer-term stock performance or the persistence of media effects on firm valuation and investment decisions. Clean technology adoption is a long-term process, and short-term market responses may not translate into sustained capital flows or technology deployment. However, the already modest explanatory power of panel regression models (low adjusted R^2) suggests that capturing long-term effects may require more complex modeling.

Fifth, the set of control variables in the model could be further expanded to include disaster data, as natural disasters may also be an important factor in explaining stock market returns of clean technology companies. An attempt was made to include natural disaster data from the EM-DAT International Disaster Database (2006-2025) (CRED, n.d.). However, this proved challenging due to the very limited number of disasters that coincided precisely with the publication dates of the news articles, as well as the heterogeneity in the units used to measure disaster magnitude (e.g., km^2 for floods, kph for tornadoes, moment magnitude for earthquakes). Including disaster type, duration, or magnitude as control variables therefore proved uninformative and even resulted in multicollinearity with sentiment and political orientation variables in the panel regression analyses in both PART 1.2 and PART 2.2.

Finally, the observed heterogeneity across clean technology sectors highlights the need for more granular, sector-specific analyses. Due to limited data availability for methanol and ammonia companies, this study could not conclusively assess their potential complementarity or substitutivity (IRENA, 2021; S&P Global, 2025). Extending both newspaper and stock market datasets beyond 2022 may increase this data coverage, enabling more robust comparative analyses of these technologies in future research.

7. Bibliography

Alam, M. S., Apergis, N., Paramati, S. R., & Fang, J. (2021). The impacts of R&D investment and stock markets on clean-energy consumption and CO emissions in OECD economies. *International Journal of Finance & Economics*, *26*(4), 4979–4992. https://doi.org/10.1002/ijfe.2049

Barber, B. M., & Odean, T. (2008). *All that Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors* (SSRN Scholarly Paper 460660). Social Science Research Network. https://doi.org/10.2139/ssrn.460660

Bettarelli, L., Furceri, D., Pizzuto, P., & Shakoor, N. (2024). Uncertainty and innovation in renewable energy. *Journal of International Money and Finance*, *149*, 103202. https://doi.org/10.1016/j.jimonfin.2024.103202

Bloomberg. (2025). Energy Transition Investment Trends 2025. *BloombergNEF*. https://about.bnef.com/energy-transition-investment/

BloombergNEF. (2025). *Global Investment in the Energy Transition Exceeded $2 Trillion for the First Time in 2024, According to BloombergNEF Report.* https://about.bnef.com/blog/global-investment-in-the-energy-transition-exceeded-2-trillion-for-the-first-time-in-2024-according-to-bloombergnef-report/

Bloomberg L.P. (n.d.). *Bloomberg Terminal* [Data service]. Bloomberg L.P.

Campbell, J. Y., Lo, A. W., & MacKinlay, A. C. (1997). *The Econometrics of Financial Markets*. Princeton University Press. https://doi.org/10.2307/j.ctt7skm5

Chadha, R., & Chaudhary, A. (2023). (PDF) A Study Analyzing an Innovative Approach to Sentiment Analysis with VADER. *ResearchGate*. https://www.researchgate.net/publication/374782701_A_Study_Analyzing_an_Innovative_Approach_to_Sentiment_Analysis_with_VADER

Chen, Y., & Mai, W. (2024). Investor attention and environmental performance of Chinese high-tech companies: The moderating effects of media attention and coverage sentiment. *Humanities and Social Sciences Communications*, *11*(1), 1–16. https://doi.org/10.1057/s41599-024-03489-1

Comin, D., & Nanda, R. (2019). *Financial Development and Technology Diffusion*. https://www.researchgate.net/publication/331337730_Financial_Development_and_Technology_Diffusion

Constantinides, G. M., Harris, M., & Stulz, R. M. (2013). *Handbook of the Economics of Finance: Corporate Finance*. Newnes.

Cornelli, G., Gambacorta, L., Oliviero, T., & Takahashi, K. (2025). *Mutual funds and climate news*. https://www.bis.org/publ/work1243.htm

CRED. (n.d.). *EM-DAT - The international disaster database*. Retrieved 25 March 2025, from https://www.emdat.be/

Cumming, D., Henriques, I., & Sadorsky, P. (2016). 'Cleantech' venture capital around the world. *International Review of Financial Analysis*, *44*, 86–97. https://doi.org/10.1016/j.irfa.2016.01.015

Da, Z., Engelberg, J., & Gao, P. (2009). *In Search of Attention* (SSRN Scholarly Paper 1364209). Social Science Research Network. https://doi.org/10.2139/ssrn.1364209

ECNO. (2024, July 2). *Clean Technologies* https://climateobservatory.eu/building-block/clean-technologies

Engelberg, J. E., & Parsons, C. A. (2011). The Causal Impact of Media in Financial Markets. *The Journal of Finance*, *66*(1), 67–97. https://doi.org/10.1111/j.1540-6261.2010.01626.x

Euronext. (2024). *Demystifying the liquidity gap between European and US equities*. https://www.euronext.com/en/news/demystifying-liquidity-gap-between-european-and-us-equities

European Commission. (n.d.). *European Climate Law*. Retrieved 21 April 2025, from https://climate.ec.europa.eu/eu-action/european-climate-law_en

European Commission. (2022). *Reducing emissions from the shipping sector.* Retrieved 9 February 2025, from https://climate.ec.europa.eu/eu-action/transport/reducing-emissions-shipping-sector_en

European Commission. (2021). *Road transport: Reducing CO_2 emissions from vehicles*. https://climate.ec.europa.eu/eu-action/transport/road-transport-reducing-co2-emissions-vehicles_en

European Commission. (2023). *The Green Deal Industrial Plan*. https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/european-green-deal/green-deal-industrial-plan_en

Fama, E. F., Fisher, L., Jensen, M. C., & Roll, R. (1969). The Adjustment of Stock Prices to New Information. *International Economic Review*, *10*(1), 1–21. https://doi.org/10.2307/2525569

Fang, L., Peress, J. (2009). *Media Coverage and the Cross-section of Stock Returns—FANG - 2009—The Journal of Finance—Wiley Online Library*. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-6261.2009.01493.x

Fang, L. H., Peress, J., & Zheng, L. (2014). *Does Media Coverage of Stocks Affect Mutual Funds' Trading and Performance?* (SSRN Scholarly Paper 1572047). Social Science Research Network. https://doi.org/10.2139/ssrn.1572047

Flammer, C. (2021). Corporate green bonds. *Journal of Financial Economics*, *142*(2), 499–516. https://doi.org/10.1016/j.jfineco.2021.01.010

Federal Reserve Bank of St. Louis. (n.d.). *FRED: Federal Reserve Economic Data*. https://fred.stlouisfed.org/

Gentzkow, M., & Shapiro, J. M. (2010). What Drives Media Slant? Evidence From U.S. Daily Newspapers. *Econometrica*, *78*(1), 35–71. https://doi.org/10.3982/ECTA7195

Ghosh, I., & Jana, R. K. (2024). Clean energy stock price forecasting and response to macroeconomic variables: A novel framework using Facebook's Prophet, NeuralProphet and explainable AI. *Technological Forecasting and Social Change*, *200*, 123148. https://doi.org/10.1016/j.techfore.2023.123148

Goldman Sachs. (2023). *The US is poised for an energy revolution*. https://www.goldmansachs.com/insights/articles/the-us-is-poised-for-an-energy-revolution

Ground News. (n.d.). Retrieved 16 February 2025, from https://ground.news/rating-system

Henderson, G. V. (1990). Problems and Solutions in Conducting Event Studies. *The Journal of Risk and Insurance*, *57*(2), 282–306. https://doi.org/10.2307/253304

Hepburn, C., Qi, Y., Stern, N., Ward, B., Xie, C., & Zenghelis, D. (2021). Towards carbon neutrality and China's 14th Five-Year Plan: Clean energy transition, sustainable urban development, and investment priorities. *Environmental Science and Ecotechnology*, *8*, 100130. https://doi.org/10.1016/j.ese.2021.100130

Hunt, L. C., Kipouros, P., & Lamprakis, Z. (2024). The Drivers of Renewable Energy: A Global Empirical Analysis of Developed and Developing Countries. *Energies*, *17*(12), Article 12. https://doi.org/10.3390/en17122902

Hunt, O. B., & Tilsted, J. P. (2024). 'Risk on steroids': Investing in the hydrogen economy. *Environment and Planning A*, 0308518X241255225. https://doi.org/10.1177/0308518X241255225

Hutto, C., & Gilbert, E. (2014). VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. *Proceedings of the International AAAI Conference on Web and Social Media*, *8*(1), Article 1. https://doi.org/10.1609/icwsm.v8i1.14550

IFP. (2023). *Barriers to Next-Gen Geothermal*. https://ifp.org/hot-rocks-part-three-barriers-to-next-gen-geothermal/

IFSWF. (2020). *Mighty oaks from little acorns grow: Sovereign wealth funds' progress on climate change*. https://www.ifswf.org/publication/mighty-oaks-little-acorns-grow-sovereign-wealth-funds%E2%80%99-progress-climate-change

IMO. (n.d.). *Future Fuels and Technology Project*. Retrieved 16 February 2025, from https://www.imo.org/en/OurWork/Environment/Pages/Future-Fuels-And-Technology.aspx

IMO. (2023). *Alternative fuel use—Regulatory status mapped*. https://www.imo.org/en/MediaCentre/Pages/WhatsNew-1841.aspx

International Energy Agency. (2018, August 1). *Statistics: World Energy Balances 2018 Overview - Event*. https://www.iea.org/events/statistics-world-energy-balances-2018-overview

International Energy Agency. (2025). *Renewable fuels – Renewables 2024 – Analysis*. Retrieved 27 April 2025, from https://www.iea.org/reports/renewables-2024/renewable-fuels

IPCC. (2022). *The evidence is clear: The time for action is now. We can halve emissions by 2030*. https://www.ipcc.ch/2022/04/04/ipcc-ar6-wgiii-pressrelease/

IRENA. (2021). *A pathway to decarbonise the shipping sector by 2050*. IRENAhttps://www.irena.org/publications/2021/Oct/A-Pathway-to-Decarbonise-the-Shipping-Sector-by-2050

Jain, H. (2024). From pollution to progress: Groundbreaking advances in clean technology unveiled. *Innovation and Green Development*, *3*(2), 100143. https://doi.org/10.1016/j.igd.2024.100143

Kaggle. (2023). *4.5M headlines from 2007-2022 [10 largest sites]*. https://www.kaggle.com/datasets/jordankrishnayah/45m-headlines-from-2007-2022-10-largest-sites

Kahneman, D. (1973). *Attention and Effort*. Prentice-Hall.

Kemp, R., & Volpi, M. (2008). The diffusion of clean technologies: A review with suggestions for future diffusion analysis. *Journal of Cleaner Production*, *16*(1, Supplement 1), S14–S21. https://doi.org/10.1016/j.jclepro.2007.10.019

Kruse, T., Mohnen, M., & Sato, M. (2024). Do Financial Markets Respond to Green Opportunities? *Journal of the Association of Environmental and Resource Economists*, *11*(3), 549–576. https://doi.org/10.1086/727370

Lanteri, A., & Rampini, A. A. (2025). *Financing the Adoption of Clean Technology* (SSRN Scholarly Paper 5159633). Social Science Research Network. https://doi.org/10.2139/ssrn.5159633

Liu, F., Kang, Y., Guo, K., & Sun, X. (2021). The relationship between air pollution, investor attention and stock prices: Evidence from new energy and polluting sectors. *Energy Policy*, *156*, 112430. https://doi.org/10.1016/j.enpol.2021.112430

Liu, L., Ge, M., & Ding, Z. (2024). Impacts of online media and investor attention on enterprise green innovation. *International Review of Economics & Finance*, *96*, 103569. https://doi.org/10.1016/j.iref.2024.103569

Loock, M. (2012). Going beyond best technology and lowest price: On renewable energy investors' preference for service-driven business models. *Energy Policy*, *40*, 21–27. https://doi.org/10.1016/j.enpol.2010.06.059

MacKinlay, A. C. (1997). Event Studies in Economics and Finance. *Journal of Economic Literature*, *35*(1), 13–39.

Maula, M. V. J., & Lukkarinen, A. (2022). *Attention Across Borders: Investor Attention as a Driver of Cross-Border Equity Crowdfunding Investments* (SSRN Scholarly Paper 3487467). Social Science Research Network. https://doi.org/10.2139/ssrn.3487467

McKinsey & Company. (2025). *Evaluating energy storage tech revenue potential*. https://www.mckinsey.com/industries/electric-power-and-natural-gas/our-insights/evaluating-the-revenue-potential-of-energy-storage-technologies

Merton, R. C. (1987). A Simple Model of Capital Market Equilibrium with Incomplete Information. *The Journal of Finance*, *42*(3), 483–510. https://doi.org/10.1111/j.1540-6261.1987.tb04565.x

Ming, Y., Li, Y., Liu, N., & Li, J. (2023). Retail investor attention and corporate environmental performance: Evidence from china. *Finance Research Letters*, *56*, 104143. https://doi.org/10.1016/j.frl.2023.104143

Muthu, S. S. (2020). 2 - Ways of measuring the environmental impact of textile processing: An overview. In S. S. Muthu (Ed.), *Assessing the Environmental Impact of Textiles and the Clothing Supply Chain (Second Edition)* (pp. 33–56). Woodhead Publishing. https://doi.org/10.1016/B978-0-12-819783-7.00002-8

Naseer, S., & St. Aubin, C. (2023, November 10). Newspapers Fact Sheet. *Pew Research Center*. https://www.pewresearch.org/journalism/fact-sheet/newspapers/

Noailly, J., Nowzohour, L., van den Heuvel, M., & Pla, I. (2024). Heard the news? Environmental policy and clean investments. *Journal of Public Economics*, *238*, 105190. https://doi.org/10.1016/j.jpubeco.2024.105190

Nyangon, J., & Darekar, A. (2024). Advancements in hydrogen energy systems: A review of levelized costs, financial incentives and technological innovations. *Innovation and Green Development*, *3*(3), 100149. https://doi.org/10.1016/j.igd.2024.100149

OECD. (2011, August 31). *The Role of Pension Funds in Financing Green Growth Initiatives*. https://www.oecd.org/en/publications/the-role-of-pension-funds-in-financing-green-growth-initiatives_5kg58j1lwdjd-en.html

OECD. (2022, August 26). *Measuring and assessing the effects of climate policy uncertainty*. https://www.oecd.org/en/publications/measuring-and-assessing-the-effects-of-climate-policy-uncertainty_34483d83-en.html

Peng, L., & Xiong, W. (2006). Investor attention, overconfidence and category learning. *Journal of Financial Economics*, *80*(3), 563–602. https://doi.org/10.1016/j.jfineco.2005.05.003

Pettengill, G. N., & Clark, J. M. (2001). Estimating Expected Returns in an Event Study Framework: Evidence from the Dartboard Column. *Quarterly Journal of Business and Economics*, *40*(3/4), 3–21.

Pham, L., & Luu Duc Huynh, T. (2020). How does investor attention influence the green bond market? *Finance Research Letters*, *35*, 101533. https://doi.org/10.1016/j.frl.2020.101533

Platform for Electromobility. (2025). *E-fuels: A costly Pandora's box for European drivers*. https://www.platformelectromobility.eu/2025/01/27/e-fuels-a-costly-pandoras-box-for-european-drivers/

Poushter, J. (2018, January 16). Many people in other countries closely follow news about the U.S. *Pew Research Center*. https://www.pewresearch.org/short-reads/2018/01/16/many-people-in-other-countries-closely-follow-news-about-the-u-s/

Reuters. (2024, October 2). *Hydrogen project investments are accelerating but uncertainty remains, IEA says*. https://www.reuters.com/sustainability/hydrogen-project-investments-are-accelerating-uncertainty-remains-iea-says-2024-10-02/

Rodrigue, J.-P. (2024). *The Geography of Transport Systems* (6th ed.). Routledge. https://doi.org/10.4324/9781003343196

Rogers, E. M. (2003). Diffusion of Innovations, 5th edition, Everett M. Rogers. Free Press, New York, NY (2003). *ResearchGate*. https://doi.org/10.1016/j.jmig.2007.07.001

Saez-Trumper, D., Castillo, C., & Lalmas, M. (2013). (PDF) Social media news communities: Gatekeeping, coverage, and statement bias. *ResearchGate*. https://doi.org/10.1145/2505515.2505623

S&P Global. (2025, March 21). *Methanol holds edge over ammonia as green marine fuel alternative in near term: Proman*. https://www.spglobal.com/commodity-insights/en/news-research/latest-news/agriculture/032125-methanol-holds-edge-over-ammonia-as-green-marine-fuel-alternative-in-near-term-proman

S&P Global Market Intelligence. (n.d.). *S&P Global Market Intelligence*. https://www.spglobal.com/market-intelligence/en

Statista. (2024). *Climate tech VC shares by region 2023*. Statista. https://www.statista.com/statistics/1197402/venture-capital-investments-in-climate-technology-startups-by-region/

Statista. (2025a). *Global energy transition investments 2024*. https://www.statista.com/statistics/1290974/investment-in-energy-transition-by-country/

Statista. (2025d). *Market size for key clean energy technologies worldwide in selected countries and regions in 2030*. https://www.statista.com/statistics/1409852/global-cleantech-market-size-country-region/

Statista. (2025e). *Most popular news websites U.S. by monthly visits 2024*. https://www.statista.com/statistics/381569/leading-news-and-media-sites-usa-by-share-of-visits/

Tetlock, P. C. (2007). Giving Content to Investor Sentiment: The Role of Media in the Stock Market. *The Journal of Finance*, *62*(3), 1139–1168. https://doi.org/10.1111/j.1540-6261.2007.01232.x

UNFCCC. (n.d.). *The Paris Agreement | UNFCCC*. Retrieved 14 April 2025, from https://unfccc.int/process-and-meetings/the-paris-agreement

United Nations. (2015). *THE 17 GOALS | Sustainable Development*. Retrieved 21 April 2025, from https://sdgs.un.org/goals

United States Congress. (1974). *H.R.16276 - A bill to further the conduct of research, development, and demonstrations in solar energy technologies, to establish Solar Energy Coordination and Management Project, to amend the National Science Foundation Act of 1950 and the National Aeronautics and Space Act of 1958, to provide for scientific and technical training in solar energy, to establish a Solar Energy Research Institute, to provide for the development of suitable incentives to assure the rapid commercial utilization of solar energy*. 93rd Congress. https://www.congress.gov/bill/93rd-congress/house-bill/16276

U.S. Department of State. (n.d.). *Climate crisis*. 2021–2025.state.gov. Retrieved May 25, 2025, from https://2021-2025.state.gov/climate-crisis/

van den Heuvel, M., & Popp, D. (2022). *The Role of Venture Capital and Governments in Clean Energy: Lessons from the First Cleantech Bubble* (SSRN Scholarly Paper 4080659). Social Science Research Network. https://papers.ssrn.com/abstract=4080659

Verdonck, C. (2025a). *NewspaperData* [Unpublished raw data].

Verdonck, C. (2025b). *StockDataFINAL2025* [Unpublished raw data].

Verdonck, C. (2025c). *SPX* [Unpublished raw data].

Verdonck, C. (2025d). *RFR* [Unpublished raw data].

Verdonck, C. (2025e). *CompanyDataFINAL* [Unpublished raw data].

Verdonck, C. (2025f). *BS_TOT_ASSETSCompanies* [Unpublished raw data].

Verdonck, C. (2025g). *NUM_OF_EMPLOYEESCompanies* [Unpublished raw data].

Verdonck, C. (2025h). *EMDATData* [Unpublished raw data].

Wilcox, R. R. (2022). Chapter 2—A Foundation for Robust Methods. In R. R. Wilcox (Ed.), *Introduction to Robust Estimation and Hypothesis Testing (Fifth Edition)* (pp. 25–44). Academic Press. https://doi.org/10.1016/B978-0-12-820098-8.00008-7

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*. The MIT Press. https://www.jstor.org/stable/j.ctt5hhcfr

World Bank (2024). *Scaling Hydrogen Financing for Development*. https://documents.worldbank.org/en/publication/documents-reports/documentdetail/en/099022024121527489

World Economic Forum. (2022, February 2). *How sustainable investing will become the norm*. https://www.weforum.org/stories/2022/02/sustainable-investing-esg-finance-future-norm/

World Economic Forum. (2024, December 12). *Here's how hard-to-abate sectors can decarbonize faster*. World Economic Forum. https://www.weforum.org/stories/2024/12/net-zero-hard-to-abate-sectors-decarbonization/

8. Appendix

Appendix 1: Screening Criteria used in S&P Capital IQ to compile list of companies in CompanyData dataset

These are the screening criteria:

1) *Industry Classifications: Electrical Equipment (Primary) OR Heavy Electrical Equipment (Primary) OR Engines and Turbines (Primary) OR Marine Transportation (Primary) OR Automobile Manufacturers (Primary) OR Hydroelectric Power Generation (Primary) OR Electric Power By Solar Energy (Primary) OR Electric Power By Wind Energy (Primary) OR Alternative Energy Resources (Primary) OR Alternative Fuels (Primary) OR Fuel Cells (Primary) OR Biofuels (Primary) OR Renewable Electricity (Primary) OR Chemicals (Primary) OR Energy (Primary) OR Machinery (Primary) OR Independent Power and Renewable Energy Producers (Primary)*
2) *Company Status: Operating*
3) *Geographic Locations: United States of America (Primary) OR United Kingdom (Primary)*
4) *Exchanges (All Listings): All Major Exchanges*

Appendix 2: Total List of Companies in CompanyData dataset

Acuity Brands, Inc.	ESS Tech, Inc.	Phoenix Motor Inc.
Advent Technologies Holdings, Inc.	Expion360 Inc.	Plug Power Inc.
AFC Energy plc	Exxon Mobil Corporation	Polar Power, Inc.
Air Products and Chemicals, Inc.	Faraday Future Intelligent Electric Inc.	PowerHouse Energy Group Plc
Alkemy Capital Investments Plc	Fluence Energy, Inc.	Preformed Line Products Company
Alternus Energy Group Plc	Flux Power Holdings, Inc.	ReNew Energy Global Plc
Altus Power, Inc.	Fly-E Group, Inc.	Rivian Automotive, Inc.
American Superconductor Corporation	FTC Solar, Inc.	Rockwell Automation, Inc.
AMETEK, Inc.	FuelCell Energy, Inc.	Sensata Technologies Holding plc
Amprius Technologies, Inc.	Gelion plc	Servotronics, Inc.
Array Technologies, Inc.	Genco Shipping & Trading Limited	SES AI Corporation
Aston Martin Lagonda Global Holdings plc	Generac Holdings Inc.	Shoals Technologies Group, Inc.
Atkore Inc.	General Motors Company	SIMEC Atlantis Energy Limited
Ayro, Inc.	GE Vernova Inc.	SKYX Platforms Corp.
Babcock & Wilcox Enterprises, Inc.	Hubbell Incorporated	Solarmax Technology Inc.
Beam Global	Ideal Power Inc.	Spruce Power Holding Corporation
Blink Charging Co.	Ilika plc	Stardust Power Inc.
Bloom Energy Corporation	Invinity Energy Systems plc	Stem, Inc.
Broadwind, Inc.	ITM Power Plc	SUNation Energy Inc.
Celanese Corporation	Kirby Corporation	Sunnova Energy International Inc.
Ceres Power Holdings plc	KULR Technology Group, Inc.	Sunrun Inc.
CF Industries Holdings, Inc.	Linde plc	Tesla, Inc.
ChargePoint Holdings, Inc.	LSB Industries, Inc.	The AES Corporation
Chevron Corporation	Luceco plc	Thermon Group Holdings, Inc.
Complete Solaria, Inc.	LyondellBasell Industries N.V.	THOR Industries, Inc.
ConnectM Technology Solutions, Inc.	Methanex Corporation	Thunder Power Holdings, Inc.
Cummins Inc.	Montauk Renewables, Inc.	Tigo Energy, Inc.
Dialight plc	Mullen Automotive, Inc.	TPI Composites, Inc.
discoverIE Group plc	NANO Nuclear Energy Inc.	Ultralife Corporation
Dragonfly Energy Holdings Corp.	NeoVolta Inc.	Vertiv Holdings Co
Drax Group plc	NET Power Inc.	Vicor Corporation
Energous Corporation	New Fortress Energy Inc.	VivoPower International PLC
Energy Focus, Inc.	Nextracker Inc.	Volex plc
Energy Vault Holdings, Inc.	Nuvve Holding Corp.	Westwater Resources, Inc.
EnerSys	nVent Electric plc	Winnebago Industries, Inc.
Enovix Corporation	Ocean Power Technologies, Inc.	Workhorse Group Inc.
Envirotech Vehicles, Inc.	Orion Energy Systems, Inc.	XPLR Infrastructure, LP
Eos Energy Enterprises, Inc.	Ormat Technologies, Inc.	Zeo Energy Corp.

Appendix 3: R Code

```
install.packages("readxl")
install.packages("dplyr")
install.packages("zoo")
install.packages("tidyr")
install.packages("plm")
install.packages("stargazer")
install.packages("lubridate")
install.packages("tseries")
install.packages("PerformanceAnalytics")
install.packages("stringr")
install.packages("data.table")
install.packages('digest')
install.packages("car")
install.packages("xtable")
install.packages("DescTools")

library(readxl)
library(ggplot2)
library(dplyr)
library(zoo)
library(tidyr)
library(plm)
library(stargazer)
library(lubridate)
library(tseries)
library(stringr)
library(data.table)
library(car)
library(xtable)
library(DescTools)

NewspaperData <- read_excel("C:/Users/clara/OneDrive - UCL/CEMS/Thesis/Thesis
R_Code/Thesis_InvestorAttention/NewspaperData.xlsx")
StockData <- read_excel("C:/Users/clara/OneDrive - UCL/CEMS/Thesis/Thesis
R_Code/Thesis_InvestorAttention/BloombergData.xlsx", sheet = "StockDataFINAL2025")
SPXData <- read_excel("C:/Users/clara/OneDrive - UCL/CEMS/Thesis/Thesis
R_Code/Thesis_InvestorAttention/BloombergData.xlsx", sheet = "SPX")
RFRData <- read_excel("C:/Users/clara/OneDrive - UCL/CEMS/Thesis/Thesis
R_Code/Thesis_InvestorAttention/RFR.xlsx", sheet = "RFR")
CompanyData <- read_excel("C:/Users/clara/OneDrive - UCL/CEMS/Thesis/Thesis
R_Code/Thesis_InvestorAttention/BloombergData.xlsx", sheet = "CompanyDataFINAL")
StockDataAssets <- read_excel("C:/Users/clara/OneDrive - UCL/CEMS/Thesis/Thesis
R_Code/Thesis_InvestorAttention/BloombergData.xlsx", sheet = "BS_TOT_ASSETSCompanies")
StockDataEmployees <- read_excel("C:/Users/clara/OneDrive - UCL/CEMS/Thesis/Thesis
R_Code/Thesis_InvestorAttention/BloombergDatas.xlsx", sheet = "NUM_OF_EMPLOYEESCompanies")
NaturalData <- read_excel("C:/Users/clara/OneDrive - UCL/CEMS/Thesis/Thesis
R_Code/Thesis_InvestorAttention/EMDATData2006-2025.xlsx", sheet = "EMDATData")

#____________________________________________________________________________

### DATA CLEANING & MANIPULATION

#____________________________________________________________________________

#Datasets to numeric
SPXData <- SPXData %>%
  mutate(across(.cols = 2, .fns = as.numeric))
StockData <- StockData %>%
  mutate(across(.cols = 2:110, .fns = as.numeric))
```

```
StockDataAssets <- StockDataAssets %>%
  mutate(across(.cols = 3:111, .fns = as.numeric))
StockDataEmployees <- StockDataEmployees %>%
  mutate(across(.cols = 3:111, .fns = as.numeric))

#Filter NaturalData for USA events in ISO column
NaturalData_USA <- NaturalData %>%
  filter(ISO == "USA") %>%
  mutate(StartDate = as.Date(StartDate)) %>%
  filter(StartDate <= as.Date("2023-12-31")) %>%
  rename('EventDate' = 'StartDate')

#Compute daily yield of RFR
RFRData <- RFRData %>%
  mutate(RFR_daily = (1 + RFR)^(1/252) - 1)

#Omit headlines who have NA sentiment
NewspaperData <- na.omit(NewspaperData)

#Code function for DailyReturn
returnfunction <- function(x) {
  returns <- (x[2:length(x)] - x[1:(length(x)-1)]) / x[1:(length(x)-1)]
  return(c(NA, returns))  # Shift down by one row
}

StockData <- StockData %>%
  arrange(Date) %>%
  mutate(
    across(
      .cols = -Date,                    # Exclude the Date column (keep all except for 1st)
      .fns  = returnfunction,           # 2) Daily return
      .names = "{.col}_Return"          # 3) New columns named "<Company>_Return"
    )
  )

##Transform StockData into StockData_long
price_cols <- colnames(StockData)[2:110]
return_cols <- colnames(StockData)[111:219]

StockData_prices_long <- StockData %>%
  pivot_longer(
    cols      = all_of(price_cols),
    names_to  = "Company",
    values_to = "ClosingPrice"
  )
StockData_prices_long <- StockData_prices_long %>%
  select(1, 111, 112)

StockData_returns_long <- StockData %>%
  pivot_longer(
    cols      = all_of(return_cols),
    names_to  = "Company",
    values_to = "DailyReturn"
  ) %>%
  # (Optional) If your return columns end in `_Return`, remove that suffix
  mutate(Company = str_remove(Company, "_Return"))
StockData_returns_long <- StockData_returns_long %>%
  select(1, 111, 112)

StockData_long <- StockData_prices_long %>%
```

```
  left_join(StockData_returns_long, by = c("Date", "Company")) %>%
  arrange(Date, Company)

head(StockData_long)

#Market Return SPX
SPXReturn_help <- returnfunction(SPXData$SPXIndex)
length(SPXReturn_help)

SPXData <- SPXData %>%
  mutate(SPX_Return_notlog = SPXReturn_help)

#Combine StockData_long & NewspaperData
PanelData_Stock_News <- NewspaperData %>%
  uncount(length(unique(StockData_long$Company)), .id = "CompanyID") %>%
  mutate(Company = unique(StockData_long$Company)[CompanyID])

PanelData_Date_Company_SPX_RFR <- StockData_long %>%
  left_join(RFRData, by = "Date") %>%
  left_join(SPXData %>% select(Date, SPX_Return_notlog), by = "Date") %>%
  left_join(CompanyData, by = "Company")

# Using na.omit to omit NA values in PanelData_Date_Company_SPX_RFR_clean
PanelData_Date_Company_SPX_RFR_clean <- na.omit(PanelData_Date_Company_SPX_RFR)

#Making columns into factors
PanelData_Stock_News$Politics <- as.factor(PanelData_Stock_News$Politics)
PanelData_Stock_News$Sentiment <- as.factor(PanelData_Stock_News$Sentiment)
PanelData_Date_Company_SPX_RFR_clean$CompanyIndustry <-
as.factor(PanelData_Date_Company_SPX_RFR_clean$CompanyIndustry)

###Pivoting StockDataAssets and StockDataEmployees
# Convert the data to long format using pivot_longer
StockDataAssets_long <- StockDataAssets %>%
  pivot_longer(cols = -c(Date, Year),          # Keep 'Date' and 'Year' columns intact
               names_to = "Company",
               values_to = "Total_Assets")

StockDataAssets_long$Year <- as.factor(StockDataAssets_long$Year)
StockDataAssets_long <- StockDataAssets_long %>%
  select(Year, Company, Total_Assets)

StockDataEmployees_long <- StockDataEmployees %>%
  pivot_longer(cols = -c(Date, Year),          # Keep 'Date' and 'Year' columns intact
               names_to = "Company",
               values_to = "Total_Employees")

StockDataEmployees_long$Year <- as.factor(StockDataEmployees_long$Year)
StockDataEmployees_long <- StockDataEmployees_long %>%
  select(Year, Company, Total_Employees)

#______________________________________________________________________________

### DESCRIPTIVE STATISTICS

#______________________________________________________________________________
###NewspaperData
# Distribution
table(NewspaperData$Newspaper)
table(NewspaperData$Sentiment)
```

```
table(NewspaperData$Politics)

# Define desired column order
desired_order <- c("Left", "Center", "Right", "Far Right")

# Reorder counts and proportions
counts <- table(NewspaperData_modified$Sentiment, NewspaperData_modified$Politics)[, desired_order]
proportions <- prop.table(counts, margin = 2)

# Combine into formatted strings: "count (proportion)"
formatted_matrix <- matrix(
  paste0(counts, " (", sprintf("%.2f", proportions), ")"),
  nrow = nrow(counts),
  dimnames = list(rownames(counts), colnames(counts))
)

# Add row totals
row_totals <- rowSums(counts)
row_total_col <- paste0(row_totals, " (", sprintf("%.2f", row_totals / sum(row_totals)), ")")
formatted_matrix_with_row_total <- cbind(formatted_matrix, Total = row_total_col)

# Add column totals
col_totals <- colSums(counts)
col_total_row <- paste0(col_totals, " (", sprintf("%.2f", col_totals / sum(col_totals)), ")")
total_total <- paste0(sum(counts), " (1.00)")
formatted_matrix_with_totals <- rbind(formatted_matrix_with_row_total, Total = c(col_total_row, total_total))

# View result
print(formatted_matrix_with_totals)

#Plots of distribution: Custom plotting function
plot_distribution <- function(data, column, title, fill_color) {
  ggplot(data, aes_string(x = column)) +
    geom_bar(fill = fill_color, alpha = 0.7, width = 0.7) +
    labs(title = title, x = column, y = "Count") +
    theme_minimal(base_size = 12) +
    theme(
      plot.title = element_text(face = "bold", hjust = 0.5, size = 14),
      axis.title = element_text(face = "bold"),
      axis.text.x = element_text(angle = 45, hjust = 1, size = 11),
      axis.text.y = element_text(size = 11)
    )
}

# Generate plots for newspaper and sentiment distribition
plot_distribution(NewspaperData, "Sentiment", "Distribution of Sentiments", "grey")


#Plot for newspaper distribution
politics_lookup <- data.frame(
  Newspaper = c("New York Times", "CNN", "FOX", "USA Today", "New York Post",
              "BBC", "Daily Mail", "CNBC", "Washington Post", "The Guardian"),
  PoliticalAlignment = c("Left", "Left", "Far Right", "Left", "Right",
                      "Center", "Right", "Center", "Left", "Left")
)

NewspaperData_colored <- NewspaperData %>%
  left_join(politics_lookup, by = "Newspaper")

ggplot(NewspaperData_colored, aes(x = Newspaper, fill = PoliticalAlignment)) +
```

```
  geom_bar(alpha = 0.8) +
  scale_fill_manual(values = c(
    "Left" = "#4A90E2",
    "Center" = "grey",
    "Right" = "#D6604D",
    "Far Right" = "#B2182B"
  )) +
  labs(title = "Distribution of Newspapers",
       x = "Newspaper", y = "Count", fill = "Political Alignment") +
  theme_minimal(base_size = 12) +
  theme(
    plot.title = element_text(face = "bold", hjust = 0.5, size = 14),
    axis.title = element_text(face = "bold"),
    axis.text.x = element_text(angle = 45, hjust = 1, size = 11),
    axis.text.y = element_text(size = 11)
  )

# Plot for politics distribution
NewspaperData_modified <- NewspaperData %>%
  mutate(
    Politics_Group = case_when(
      Politics %in% c("Far Right", "Right") ~ "Right",
      TRUE ~ Politics
    )
  )

NewspaperData_modified <- NewspaperData_modified %>%
  mutate(Politics_Detail = Politics)

ggplot(NewspaperData_modified, aes(x = Politics_Group, fill = Politics_Detail)) +
  geom_bar(width = 0.7, alpha = 0.8) +
  scale_fill_manual(values = c("Left" = "#4A90E2", "Center" = "grey",
                               "Right" = "#D6604D", "Far Right" = "#B2182B")) +
  labs(
    title = "Distribution of Political Alignment",
    x = "Political Alignment",
    y = "Count",
    fill = "Category"
  ) +
  theme_minimal(base_size = 12) +
  theme(
    plot.title = element_text(face = "bold", hjust = 0.5, size = 14),
    axis.title = element_text(face = "bold"),
    axis.text = element_text(size = 11),
    legend.position = "right"
  )


###Text Analysis
keywords <- c("clean tech", "clean technology", "green technology", "renewable energy", "green power", "solar power", "solar panel", "wind power", "wind turbine", "wind farm", "hydropower", "bioenergy", "biofuel", "geothermal power", "electric vehicle", "battery storage", "energy storage", "lithium battery", "smart grid", "hydrogen", "methanol", "ammonia", "hydroelectric", "alternative fuels", "fuel cells", "carbon capture", "energy efficiency", "wave energy", "biogas","net zero emissions", "sustainable development", "energy transition", "green building", "smart thermostats", "energy management systems")

# Function to count occurrences of each keyword
count_keywords <- function(headlines, keywords) {
  keyword_counts <- setNames(vector("integer", length(keywords)), keywords)
  for (headline in headlines) {
```

```
    # Make sure to handle case insensitivity and full words
    words <- tolower(headline)
    for (keyword in keywords) {
      keyword_pattern <- paste0("\\b", keyword, "\\b") # regex to match whole words only
      keyword_counts[keyword] <- keyword_counts[keyword] + sum(grepl(keyword_pattern, words, perl =
TRUE))
    }
  }
  return(keyword_counts)
}

# Calculate keyword occurrences
technology_counts <- count_keywords(NewspaperData$Headline, keywords)

print(technology_counts)
sum(technology_counts)

# Convert to data frame for plotting
technology_df <- as.data.frame(technology_counts)
technology_df$Technology <- rownames(technology_df)
rownames(technology_df) <- NULL
colnames(technology_df) <- c("Count", "Technology")

technology_df_filtered <- technology_df %>%
  filter(Count > 45) %>%
  arrange(desc(Count))  # optional: sort by frequency

# Plot the results
ggplot(technology_df, aes(x = Technology, y = Count)) +
  geom_bar(stat = "identity", fill = "grey", alpha = 0.8) +
  theme_classic(base_size = 12) +
  theme(
    axis.text.x = element_text(angle = 45, hjust = 1),
    plot.title = element_text(face = "bold", hjust = 0.5, size = 14),
    axis.title = element_text(face = "bold")
  ) +
  labs(
    title = "Distribution of Technologies Mentioned in Headlines",
    x = "Technology",
    y = "Count"
  )

ggplot(technology_df_filtered, aes(x = reorder(Technology, Count), y = Count)) +
  geom_bar(stat = "identity", fill = "grey", alpha = 0.8) +
  coord_flip() +
  labs(
    title = "Most Frequently Mentioned Clean Technologies in Headlines",
    x = "Technology Keyword",
    y = "Count"
  ) +
  theme_classic(base_size = 12) +
  theme(
    plot.title = element_text(face = "bold", hjust = 0.5, size = 14),
    axis.title = element_text(face = "bold"),
    axis.text = element_text(size = 11)
  )

###CompanyData
#Convert relevant columns to factors
CompanyData$CompanyIndustry <- as.factor(CompanyData$CompanyIndustry)
```

```
CompanyData$CompanyCountry <- as.factor(CompanyData$CompanyCountry)


#Distribution
table(CompanyData$CompanyCountry)
table(CompanyData$CompanyIndustry)
table(CompanyData$CompanyAmmonia)
table(CompanyData$CompanyMethanol)
table(CompanyData$CompanyHydrogen)
table(CompanyData$CompanyEfuel)

#Plots
plot_distribution(CompanyData, "CompanyCountry", "Distribution of U.S. and UK Companies", "grey")


# Plot for CompanyIndustry
plot_distribution(CompanyData, "CompanyIndustry", "Distribution of Company Industries", "grey")



#______________________________________________________________________________

###ANALYSIS: PART 1
###PART 1.1 EVENT STUDY

#______________________________________________________________________________


###Event Study for 1 EventDate "2021-07-02"
##alpha and beta for each company
# Define event date
event_date <- as.Date("2021-07-02")

# Define estimation window
estimation_start <- event_date - 220
estimation_end   <- event_date - 21

# Filter estimation window data (subset dataset to only include estimation window)
estimation_data <- PanelData_Date_Company_SPX_RFR_clean %>%
  filter(Date >= estimation_start & Date <= estimation_end)

# Get list of companies (all unique companies in estimation_data)
company_list <- unique(estimation_data$Company)

# Create empty dataframe to store alpha and beta OLS coefficients
alpha_beta_matrix <- data.frame(
  Company = company_list,
  Alpha = NA,
  Beta = NA
)

# Loop through each company and estimate alpha and beta via OLS
for (i in seq_along(company_list)) {
  company_i <- company_list[i]

  df_i <- estimation_data %>%
    filter(Company == company_i)

  if (nrow(df_i) >= 30) {  # Ensure sufficient data points
    model <- lm(DailyReturn ~ SPX_Return_notlog, data = df_i)
```

```
    alpha_beta_matrix$Alpha[i] <- coef(model)[1]
    alpha_beta_matrix$Beta[i] <- coef(model)[2]
  }
}

###CAR with market model and CAPM for one EventDate "2021-07-02"
#Define big event window
big_window <- -20:60


# Subset the panel for that entire range
# Then merge alpha-beta estimates
# Finally, compute 'Day'
event_data <- PanelData_Date_Company_SPX_RFR_clean %>%
  filter(Date %in% (event_date + big_window)) %>%
  left_join(alpha_beta_matrix, by = "Company") %>%
  mutate(
    Day = as.integer(as.Date(Date) - as.Date(event_date)),
    # Market model AR
    ExpectedReturn = Alpha + Beta * SPX_Return_notlog,
    AR_market = DailyReturn - ExpectedReturn,
    #CAPM model AR
    ExcessReturn = RFR_daily + Beta * (SPX_Return_notlog - RFR_daily),
    AR_capm = DailyReturn - ExcessReturn
  )


#Define significance stars
get_significance <- function(p) {
  if (is.na(p)) return(NA)
  if (p < 0.01) return("***")
  else if (p < 0.05) return("**")
  else if (p < 0.1) return("*")
  else return("")
}

# Named list of windows: c(start_day, end_day)
event_windows <- list(
  "[-5,  0]" = c(-5, 0),
  "[-5,   10]" = c(-5, 10),
  "[0,    10]" = c(0, 10),
  "[11,   20]" = c(11, 20)
)

CAR_window_results <- data.frame()

for (window_name in names(event_windows)) {
  range_vec <- event_windows[[window_name]]
  start_day <- range_vec[1]
  end_day   <- range_vec[2]

  # 1) Subset to days in [start_day, end_day]
  window_data <- event_data %>%
    filter(Day >= start_day & Day <= end_day)

  # 2) Sum AR for each company
  CAR_df <- window_data %>%
    group_by(Company) %>%
    summarise(
      CAR_market = sum(AR_market, na.rm = TRUE),
```

```
    CAR_capm = sum(AR_capm, na.rm = TRUE),
    .groups = "drop",
   )

  # 3) T-test the CAR distribution
  test_market <- t.test(CAR_df$CAR_market)
  test_capm <- t.test(CAR_df$CAR_capm)

  # 4) Store results in a row
  window_summary <- data.frame(
   Window = window_name,
   StartDay = start_day,
   EndDay   = end_day,
   #Market Model results
   MeanCAR_market  = mean(CAR_df$CAR_market, na.rm = TRUE),
   t_stat_market   = test_market$statistic,
   p_value_market  = test_market$p.value,
   Significance_market = get_significance(test_market$p.value),
   #Capm Model results
   MeanCAR_capm = mean(CAR_df$CAR_capm, na.rm = TRUE),
   t_stat_capm  = test_capm$statistic,
   p_value_capm = test_capm$p.value,
   Significance_capm = get_significance(test_capm$p.value)
  )

  # 5) Append to master results data frame
  CAR_window_results <- rbind(CAR_window_results, window_summary)
}

#____________________________________________________________

### Same thing for different event_date
event_date2 <- as.Date("2022-03-01")

# Define estimation window
estimation_start2 <- event_date2 - 220
estimation_end2   <- event_date2 - 21

# Filter estimation window data
estimation_data2 <- PanelData_Date_Company_SPX_RFR_clean %>%
  filter(Date >= estimation_start2 & Date <= estimation_end2)

# Get list of companies
company_list2 <- unique(estimation_data2$Company)

# Create empty dataframe to store alpha and beta OLS coefficients
alpha_beta_matrix2 <- data.frame(
  Company = company_list2,
  Alpha = NA,
  Beta = NA
)

# Loop through each company and estimate alpha and beta via OLS
for (i in seq_along(company_list2)) {
  comp_i <- company_list2[i]

  df_i <- estimation_data2 %>%
    filter(Company == comp_i)

  if (nrow(df_i) >= 30) {
```

```
    model <- lm(DailyReturn ~ SPX_Return_notlog, data = df_i)
    alpha_beta_matrix2$Alpha[i] <- coef(model)[1]
    alpha_beta_matrix2$Beta[i]  <- coef(model)[2]
  }
}

###CAR for one event_date "2022-03-01"
# Big window
big_window2 <- -20:60

# Subset, merge alpha-beta, compute 'Day'
event_data2 <- PanelData_Date_Company_SPX_RFR_clean %>%
  filter(Date %in% (event_date2 + big_window2)) %>%
  left_join(alpha_beta_matrix2, by = "Company") %>%
  mutate(
    Day = as.integer(as.Date(Date) - as.Date(event_date2)),
    # Market model AR
    ExpectedReturn = Alpha + Beta * SPX_Return_notlog,
    AR_market = DailyReturn - ExpectedReturn,
    # CAPM-based AR
    ExcessReturn = RFR_daily + Beta * (SPX_Return_notlog - RFR_daily),
    AR_capm = DailyReturn - ExcessReturn
  )

#Compute CARs for different windows using market and CAPM model
CAR_window_results2 <- data.frame()

for (window_name in names(event_windows)) {
  range_vec <- event_windows[[window_name]]
  start_day <- range_vec[1]
  end_day   <- range_vec[2]

  # Subset event_data2 for this window
  window_data2 <- event_data2 %>%
    filter(Day >= start_day & Day <= end_day)

  # Sum AR for each company
  CAR_df2 <- window_data2 %>%
    group_by(Company) %>%
    summarise(
      CAR_market = sum(AR_market, na.rm = TRUE),
      CAR_capm   = sum(AR_capm,   na.rm = TRUE),
      .groups = "drop"
    )

  # Conduct t-tests
  test_market2 <- t.test(CAR_df2$CAR_market)
  test_capm2   <- t.test(CAR_df2$CAR_capm)

  # Build summary row
  window_summary2 <- data.frame(
    Window     = window_name,
    StartDay   = start_day,
    EndDay     = end_day,

    # Market Model
    MeanCAR_market  = mean(CAR_df2$CAR_market, na.rm = TRUE),
    t_stat_market   = test_market2$statistic,
    p_value_market  = test_market2$p.value,
    Significance_market = get_significance(test_market$p.value),
```

```
  # CAPM Model
  MeanCAR_capm    = mean(CAR_df2$CAR_capm, na.rm = TRUE),
  t_stat_capm     = test_capm2$statistic,
  p_value_capm    = test_capm2$p.value,
  Significance_capm = get_significance(test_capm$p.value)
 )

 # Append to final results
 CAR_window_results2 <- rbind(CAR_window_results2, window_summary2)
}

#______________________________________________________________

###EventStudy for each EventDate for each company: CAR
# 1. Define event windows --> the same as for single event trial

# 2. Create an empty dataframe to store all CARs from all subwindows and events
CAR_all_windows <- data.frame()

# 3. Loop over each event in NewspaperData
for (i in 1:nrow(NewspaperData)) {

 # Extract the event date and headline
 event_date <- as.Date(NewspaperData$EventDate[i])
 headline_text <- NewspaperData$Headline[i]

 # --- A) Compute alpha and beta from estimation window [-220, -21] ---
 estimation_start <- event_date - 220
 estimation_end   <- event_date - 21

 estimation_data <- PanelData_Date_Company_SPX_RFR_clean %>%
  filter(Date >= estimation_start & Date <= estimation_end)

 # Unique companies in estimation window
 company_list <- unique(estimation_data$Company)

 # Estimate alpha-beta
 alpha_beta_list <- lapply(company_list, function(comp) {
  firm_data <- estimation_data %>%
   filter(Company == comp)

  if (nrow(firm_data) >= 20) {
   model <- lm(DailyReturn ~ SPX_Return_notlog, data = firm_data)
   data.frame(
    Company = comp,
    Alpha   = coef(model)[1],
    Beta    = coef(model)[2]
   )
  } else {
   data.frame(
    Company = comp,
    Alpha   = NA,
    Beta    = NA
   )
  }
 })
 alpha_beta_df <- dplyr::bind_rows(alpha_beta_list)

 # If no valid alpha-beta, skip
```

```
  if (nrow(alpha_beta_df) == 0 || !"Company" %in% names(alpha_beta_df)) {
    cat("Skipping event on", event_date, "due to missing alpha/beta estimates\n")
    next
  }

  # --- B) Build big window [-20, +60], compute AR for each day ---
  big_window <- -20:60
  event_data <- PanelData_Date_Company_SPX_RFR_clean %>%
    filter(Date %in% (event_date + big_window)) %>%
    left_join(alpha_beta_df, by = "Company") %>%
    mutate(
      Day = as.integer(as.Date(Date) - event_date),
      # ExcessReturn (CAPM approach)
      ExcessReturn = RFR_daily + Beta * (SPX_Return_notlog - RFR_daily),
      AR = DailyReturn - ExcessReturn
    )

  # --- C) For each subwindow, compute CAR and store ---
  for (window_name in names(event_windows)) {
    range_vec <- event_windows[[window_name]]
    start_day <- range_vec[1]
    end_day   <- range_vec[2]

    window_data <- event_data %>%
      filter(Day >= start_day & Day <= end_day)

    # Sum AR per company
    CAR_df <- window_data %>%
      group_by(Company) %>%
      summarise(CAR = sum(AR, na.rm = TRUE), .groups = "drop")

    # Combine into a single table
    # Each row = (event_date, window_name, company, CAR)
    if (nrow(CAR_df) > 0) {
      CAR_window <- CAR_df %>%
        mutate(
          EventDate = event_date,
          Headline  = headline_text,
          Window    = window_name,
          StartDay  = start_day,
          EndDay    = end_day
        )

      # Append to master
      CAR_all_windows <- dplyr::bind_rows(CAR_all_windows, CAR_window)
    }
  }

  # Progress monitor (optional)
  if (i %% 25 == 0) {
    cat("Processed", i, "of", nrow(NewspaperData), "events\n")
  }
}

# the company NasdaqCM:SKYX CARs and NASDAQ:CM:BLNK have extreme values (from -1000 to +1000
and up until 65), so I filter it out to not bias the t-tests
CAR_all_windows <- CAR_all_windows %>%
  filter(Company != "NasdaqCM:SKYX")

upper_threshold <- quantile(CAR_all_windows$CAR, probs = 0.99995, na.rm = TRUE)  #0.99995
```

```
CAR_all_windows <- CAR_all_windows %>%
  mutate(CAR_winsor = ifelse(CAR > upper_threshold, upper_threshold, CAR))

# T-test for each event & window
CAR_event_significance <- CAR_all_windows %>%
  group_by(EventDate, Headline, Window) %>%
  summarise(
    Mean_CAR = mean(CAR, na.rm = TRUE),
    t_stat   = t.test(CAR)$statistic,
    p_value  = t.test(CAR)$p.value,
    Significance = get_significance(p_value),
    .groups = "drop"
  )

#T-test on mean CAR per company
CAR_company_significance <- CAR_all_windows %>%
  group_by(Company, Window) %>%
  summarise(
    Mean_CAR = mean(CAR, na.rm = TRUE),
    t_stat   = t.test(CAR)$statistic,
    p_value  = t.test(CAR)$p.value,
    Significance = get_significance(p_value),
    .groups = "drop"
  )

#T-test on mean CAR_winsor per company
CAR_company_significance_winsor <- CAR_all_windows %>%
  group_by(Company, Window) %>%
  summarise(
    Mean_CAR_winsor = mean(CAR_winsor, na.rm = TRUE),
    t_stat   = t.test(CAR_winsor)$statistic,
    p_value  = t.test(CAR_winsor)$p.value,
    Significance = get_significance(p_value),
    .groups = "drop"
  )

#T-test on mean CAR per event window
CAR_subwindow_significance <- CAR_all_windows %>%
  group_by(Window) %>%
  summarise(
    Mean_CAR = mean(CAR, na.rm = TRUE),
    t_stat   = t.test(CAR)$statistic,
    p_value  = t.test(CAR)$p.value,
    Significance = get_significance(p_value),
    .groups = "drop"
  )

CAR_subwindow_significance_winsor <- CAR_all_windows %>%
  group_by(Window) %>%
  summarise(
    Mean_CAR_winsor = mean(CAR_winsor, na.rm = TRUE),
    t_stat   = t.test(CAR_winsor)$statistic,
    p_value  = t.test(CAR_winsor)$p.value,
    Significance = get_significance(p_value),
    .groups = "drop"
  )

latex_table <- xtable(CAR_subwindow_significance)

# Binomial test per event window: do most companies have CAR > 0?
```

```
CAR_directional_test <- CAR_company_significance %>%
  filter(!is.na(Mean_CAR)) %>%
  group_by(Window) %>%
  summarise(
    n_total = n(),
    n_positive = sum(Mean_CAR > 0),
    proportion_positive = n_positive / n_total,
    binom_p_value = binom.test(n_positive, n_total, p = 0.5, alternative = "two.sided")$p.value,
    Direction = case_when(
      proportion_positive > 0.5 ~ "Mostly Positive",
      proportion_positive < 0.5 ~ "Mostly Negative",
      TRUE ~ "Even"
    ),
    .groups = "drop"
  )

# Binomial test per event window: do most companies have CAR_winsor > 0? --> does not change for winsor compared to regular CAR
CAR_directional_test_winsor <- CAR_company_significance_winsor %>%
  filter(!is.na(Mean_CAR_winsor)) %>%
  group_by(Window) %>%
  summarise(
    n_total = n(),
    n_positive = sum(Mean_CAR_winsor > 0),
    proportion_positive = n_positive / n_total,
    binom_p_value = binom.test(n_positive, n_total, p = 0.5, alternative = "two.sided")$p.value,
    Direction = case_when(
      proportion_positive > 0.5 ~ "Mostly Positive",
      proportion_positive < 0.5 ~ "Mostly Negative",
      TRUE ~ "Even"
    ),
    .groups = "drop"
  )


###Focusing on the [0,10] subset: for CAR and CAR_winsor
CAR_0_10 <- CAR_all_windows %>%
  filter(Window == "[0,    10]") %>%
  rename(CAR_0_10 = CAR)

#____________________________________________________________________________

###ANALYSIS PART 2.2 REGRESSION

#____________________________________________________________________________
###Merge CAR_0_10 with PanelData_Stock_News
PanelData_Stock_News <- PanelData_Stock_News %>%
  left_join(CAR_0_10 %>% select(Company, EventDate, Headline, CAR_0_10, CAR_winsor),
            by = c("Company", "EventDate", "Headline")
  )

### Regression Analysis
PanelData_Stock_News <- PanelData_Stock_News %>%
  mutate(
    Year = factor(year(EventDate)),
    MonthYear = format(EventDate, "%Y-%m")
  ) %>%
  na.omit()

#Reorder factor levels so "Center" and "neutral" is baseline
```

```
PanelData_Stock_News$Politics <- relevel(PanelData_Stock_News$Politics, ref = "Center")
PanelData_Stock_News$Sentiment <- relevel(PanelData_Stock_News$Sentiment, ref = "Neutral")

#Add StockDataAssets and StockDataEmployees to PanelData_Stock_News
PanelData_Stock_News <- PanelData_Stock_News %>%
  left_join(StockDataAssets_long, by = c("Year", "Company"))
PanelData_Stock_News <- PanelData_Stock_News %>%
  left_join(StockDataEmployees_long, by = c("Year", "Company"))

#df that will be used: PanelData_Stock_News + PanelData_Date_Company_SPX_RFR_clean
df <- PanelData_Stock_News %>%
  left_join(
    PanelData_Date_Company_SPX_RFR_clean %>%
      select(Company, CompanyCountry, CompanyName, USCompany, CompanyIndustry, CompanyEV,
CompanyAmmonia, CompanyMethanol, CompanyHydrogen, CompanyEfuel) %>%
      distinct(),
    by = "Company"
  )

#Add NaturalData_USA data to df
NaturalData_USA <- NaturalData_USA %>%
  mutate(EventDate = as.Date(EventDate))

NaturalData_USA_dedup <- NaturalData_USA %>%
  distinct(EventDate, .keep_all = TRUE)

df <- df %>%
  mutate(
    EventDate = as.Date(as.character(EventDate)))  # convert factor to character to date

df <- df %>%
  left_join(NaturalData_USA_dedup, by = "EventDate")

df <- df %>%
  mutate(
    DisasterSubgroup = as.factor(DisasterSubgroup),
    DisasterType = as.factor(DisasterType),
    MonthYear = as.factor(MonthYear)
  )

df <- df %>%
  mutate(Total_Assets_log = log(Total_Assets + 1))
df <- df %>%
  mutate(Total_Employees_log = log(Total_Employees))


#Account for fixed effects -> transform into factors
df$EventDate <- factor(df$EventDate)
df$Company <- as.factor(df$Company)
df$Sentiment <- as.factor(df$Sentiment)
df$Newspaper   <- as.factor(df$Newspaper)
df$NewsCountry <- as.factor(df$NewsCountry)
df$CompanyIndustry <- as.factor(df$CompanyIndustry)
df$StartYear <- as.factor(df$StartYear)
df$USCompany <- as.factor(df$USCompany)

#Create extra binary columns when interaction is 1
df <- df %>%
  mutate(
    MatchEV      = if_else(NewsEV == 1 & CompanyEV == 1, 1, 0),
```

```
    MatchAmmonia  = if_else(NewsAmmonia == 1 & CompanyAmmonia == 1, 1, 0),
    MatchMethanol = if_else(NewsMethanol == 1 & CompanyMethanol == 1, 1, 0),
    MatchHydrogen = if_else(NewsHydrogen == 1 & CompanyHydrogen == 1, 1, 0),
    MatchEfuel    = if_else(NewsEfuel == 1 & CompanyEfuel == 1, 1, 0)
  )

#panel regression for hypothesis 2: tone effect
plm1 <- plm(
  CAR_winsor ~ Sentiment + NewsCountry + USCompany + CompanyIndustry +
    Total_Assets_log + Total_Employees_log + Year,
  data = df,
  index = c("Company", "EventDate", "Headline"),
  model = "within",          # "within" = fixed effects
  effect = "individual"      # or "twoways" if you want time FE too
)
summary(plm1)

vector = table(index(df), useNA = "ifany")
vector[2]

lm1 <- lm(
  CAR_winsor ~ Sentiment +
    NewsCountry + USCompany + CompanyIndustry + Total_Assets_log +
    Total_Employees_log + Year,
  data = df
)
summary(lm1)

vif(lm1, type = "predictor")
alias(lm1)

#panel regression for hypothesis 3: politics effect
plm2 <- plm(
  CAR_winsor ~ Politics + NewsCountry + USCompany + CompanyIndustry +
    Total_Assets_log + Total_Employees_log + Year,
  data = df,
  index = c("Company", "EventDate", "Headline"),
  model = "within",          # "within" = fixed effects
  effect = "individual"      # or "twoways" if you want time FE too
)
summary(plm2)

lm2 <- lm(
  CAR_winsor ~ Politics +
    NewsCountry + USCompany + CompanyIndustry + Total_Assets_log +
    Total_Employees_log + Year,
  data = df
)
summary(lm2)

vif(lm2, type = "predictor")
alias(lm2)


#panel regression for hypothesis 4: Interaction effect tone & politics
plm3 <- plm(
  CAR_winsor ~ Sentiment * Politics +
    NewsCountry + USCompany + CompanyIndustry +
    Total_Assets_log + Total_Employees_log + Year,
  data = df,
```

```
  index = c("Company", "EventDate", "Headline"),
  model = "within",          # "within" = fixed effects
  effect = "individual"
)
summary(plm3)

lm3 <- lm(
  CAR_winsor ~ Sentiment * Politics +
    NewsCountry + USCompany + CompanyIndustry + Total_Assets_log +
    Total_Employees_log + Year,
  data = df
)
summary(lm3)

vif(lm3, type = "predictor")
alias(lm3)


#panel regression with more controls for company size (fixed effects)
#No Newspaper, DisasterType, DisasterSubgroup
plm4 <- plm(
  CAR_winsor ~ Sentiment + Politics + Sentiment * Politics +
    NewsCountry + USCompany + CompanyIndustry + Total_Assets_log +
    Total_Employees_log + Year,
  data = df,
  index = c("Company", "EventDate", "Headline"),
  model = "within",
  effect = "individual"
)
summary(plm4)

lm4 <- lm(
  CAR_winsor ~ Sentiment + Politics + Sentiment * Politics +
    NewsCountry + USCompany + CompanyIndustry + Total_Assets_log +
    Total_Employees_log + Year + DisasterLength,
  data = df
)
summary(lm4)

vif(lm4, type = "predictor")
alias(lm4)


#Verify with CAR_0_10 (not winsor)
plm5 <- plm(
  CAR_0_10 ~ Sentiment * Politics +
    NewsCountry + USCompany + CompanyIndustry +
    Total_Assets_log + Total_Employees_log + Year,
  data = df,
  index = c("Company", "EventDate", "Headline"),
  model = "within",          # "within" = fixed effects
  effect = "individual"      # or "twoways" if you want time FE too
)
summary(plm5)

lm5 <- lm(
  CAR_0_10 ~ Sentiment * Politics +
    NewsCountry + USCompany + CompanyIndustry + Total_Assets_log +
    Total_Employees_log + Year,
  data = df
```

```
)
summary(lm5)

vif(lm5, type = "predictor")
alias(lm5)


# Create the regression table
stargazer(plm1, plm2, plm3,
        type = "latex",   # use "latex" for LaTeX output and "text" for text output
        title = "Panel Regression Results",
        column.labels = c("Tone Effect", "Politics Effect", "Interaction Effect", "With Company Controls"),
        dep.var.labels = "CAR_0_10_winsor",
        model.names = FALSE,
        omit.stat = c("f", "ser"),
        no.space = TRUE,
        digits = 3)

stargazer(plm4,
        type = "latex",   # use "latex" for LaTeX output and "text" for text output
        title = "Panel Regression Results",
        column.labels = c("Tone Effect", "Politics Effect", "Interaction Effect", "With Company Controls"),
        dep.var.labels = "CAR_0_10_winsor",
        model.names = FALSE,
        omit.stat = c("f", "ser"),
        no.space = TRUE,
        digits = 3)

#______________________________________________________________________________

#PART 2 SUBGROUP COMPARISON
###ANALYSIS PART 2.1 EVENT STUDY by Subgroup Alignment

#______________________________________________________________________________

###Matrix
# Define news and company variables
news_vars <- c("NewsAmmonia", "NewsMethanol", "NewsHydrogen", "NewsEfuel", "NewsEV")
non_news_vars <- paste0("NewsNo", c("Ammonia", "Methanol", "Hydrogen", "Efuel", "EV"))
company_vars <- c("CompanyAmmonia", "CompanyMethanol", "CompanyHydrogen", "CompanyEfuel",
"CompanyEV")

# Add non-news variables to the dataset
df <- df %>%
  mutate(
    NewsNoAmmonia = ifelse(NewsAmmonia == 0, 1, 0),
    NewsNoMethanol = ifelse(NewsMethanol == 0, 1, 0),
    NewsNoHydrogen = ifelse(NewsHydrogen == 0, 1, 0),
    NewsNoEfuel = ifelse(NewsEfuel == 0, 1, 0),
    NewsNoEV = ifelse(NewsEV == 0, 1, 0)
  )

# Initialize matrix
row_labels <- c(news_vars, non_news_vars)
results_matrix <- matrix("", nrow = length(row_labels), ncol = length(company_vars))
rownames(results_matrix) <- row_labels
colnames(results_matrix) <- company_vars

# Loop through each news and company combination
for (i in seq_along(row_labels)) {
```

```
  for (j in seq_along(company_vars)) {
    subset_df <- df %>%
      filter(.data[[row_labels[i]]] == 1, .data[[company_vars[j]]] == 1)

    mean_car <- mean(subset_df$CAR_winsor, na.rm = TRUE)
    t_test <- t.test(subset_df$CAR_winsor, mu = 0)
    star <- get_significance(t_test$p.value)

    # Round to 4 digits and append star
    results_matrix[i, j] <- paste0(round(mean_car, 4), star)
  }
}

# Convert to data frame for visualization
results_df <- as.data.frame(results_matrix)
results_df <- cbind(NewsType = row_labels, results_df)
print(results_df)


#________________________________________________________________________________

###ANALYSIS PART 2.2 REGRESSION with Technology-Topic Interaction

#________________________________________________________________________________

###Making the different subsets in the df dataset per news-company alignment
NewsHydrogenCompanyHydrogen <- df %>%
  filter(NewsHydrogen == 1 & CompanyHydrogen == 1)
NewsAmmoniaCompanyAmmonia <- df %>%
  filter(NewsAmmonia == 1 & CompanyAmmonia == 1)
NewsMethanolCompanyMethanol <- df %>%
  filter(NewsMethanol == 1 & CompanyMethanol == 1)
NewsEfuelCompanyEfuel <- df %>%
  filter(NewsEfuel == 1 & CompanyEfuel == 1)
NewsEVCompanyEV <- df %>%
  filter(NewsEV == 1 & CompanyEV == 1)


#Panel Regressions:
#Hydrogen: No Newspaper, DisasterType, DisasterSubgroup
plm6_sent <- plm(
  CAR_winsor ~ Sentiment +
    NewsCountry + USCompany + CompanyIndustry +
    Total_Assets_log + Total_Employees_log + Year,
  data = NewsHydrogenCompanyHydrogen,
  index = c("Company", "EventDate", "Headline"),
  model = "within",          # "within" = fixed effects
  effect = "individual"      # or "twoways" if you want time FE too
)
summary(plm6_sent)

plm6_pol <- plm(
  CAR_winsor ~ Politics +
  NewsCountry + USCompany + CompanyIndustry +
    Total_Assets_log + Total_Employees_log + Year,
  data = NewsHydrogenCompanyHydrogen,
  index = c("Company", "EventDate", "Headline"),
  model = "within",          # "within" = fixed effects
  effect = "individual"      # or "twoways" if you want time FE too
)
```

```r
summary(plm6_pol)


plm6 <- plm(
  CAR_winsor ~ Sentiment + Politics + Sentiment * Politics +
    NewsCountry + USCompany + CompanyIndustry +
    Total_Assets_log + Total_Employees_log + Year,
  data = NewsHydrogenCompanyHydrogen,
  index = c("Company", "EventDate", "Headline"),
  model = "within",          # "within" = fixed effects
  effect = "individual"      # or "twoways" if you want time FE too
)
summary(plm6)

lm6 <- lm(
  CAR_winsor ~ Sentiment * Politics +
    NewsCountry + USCompany + CompanyIndustry + Total_Assets_log +
    Total_Employees_log + Year,
  data = NewsHydrogenCompanyHydrogen
)
summary(lm6)

vif(lm6, type = "predictor")
alias(lm6)


plm11 <- plm(
  CAR_0_10 ~ Sentiment + Politics + Sentiment * Politics +
    NewsCountry + USCompany + CompanyIndustry +
    Total_Assets_log + Total_Employees_log + Year,
  data = NewsHydrogenCompanyHydrogen,
  index = c("Company", "EventDate", "Headline"),
  model = "within",          # "within" = fixed effects
  effect = "individual"      # or "twoways" if you want time FE too
)
summary(plm11)

#Ammonia: No Newspaper, DisasterType, DisasterSubgroup, Sentiment * Politics, Year
plm7 <- plm(
  CAR_winsor ~ Sentiment + Politics +
    NewsCountry + USCompany + CompanyIndustry +
    Total_Assets_log + Total_Employees_log,
  data = NewsAmmoniaCompanyAmmonia,
  index = c("Company", "EventDate", "Headline"),
  model = "within",          # "within" = fixed effects
  effect = "individual"      # or "twoways" if you want time FE too
)
summary(plm7)

lm7 <- lm(
  CAR_winsor ~ Sentiment + Politics +
    NewsCountry + USCompany + CompanyIndustry +
    Total_Assets_log + Total_Employees_log,
  data = NewsAmmoniaCompanyAmmonia
)
summary(lm7)

vif(lm7, type = "predictor")
alias(lm7)
```

```
plm12 <- plm(
  CAR_0_10 ~ Sentiment + Politics +
    NewsCountry + USCompany + CompanyIndustry +
    Total_Assets_log + Total_Employees_log,
  data = NewsAmmoniaCompanyAmmonia,
  index = c("Company", "EventDate", "Headline"),
  model = "within",          # "within" = fixed effects
  effect = "individual"      # or "twoways" if you want time FE too
)
summary(plm12)

#Methanol: No Newspaper, DisasterType, DisasterSubgroup, Year, Sentiment * Politics
plm8 <- plm(
  CAR_winsor ~ Sentiment + Politics +
    NewsCountry + USCompany + CompanyIndustry +
    Total_Assets_log + Total_Employees_log,
  data = NewsMethanolCompanyMethanol,
  index = c("Company", "EventDate", "Headline"),
  model = "within",          # "within" = fixed effects
  effect = "individual"      # or "twoways" if you want time FE too
)
summary(plm8)

lm8 <- lm(
  CAR_winsor ~ Sentiment + Politics +
    NewsCountry + USCompany + CompanyIndustry +
    Total_Assets_log + Total_Employees_log,
  data = NewsMethanolCompanyMethanol
)
summary(lm8)

vif(lm8, type = "predictor")
alias(lm8)

#Efuel: No Newspaper, DisasterType, DisasterSubgroup, Year, Sentiment * Politics
plm9 <- plm(
  CAR_winsor ~ Sentiment + Politics +
    NewsCountry + USCompany + CompanyIndustry +
    Total_Assets_log + Total_Employees_log,
  data = NewsEfuelCompanyEfuel,
  index = c("Company", "EventDate", "Headline"),
  model = "within",          # "within" = fixed effects
  effect = "individual"      # or "twoways" if you want time FE too
)
summary(plm9)

lm9 <- lm(
  CAR_winsor ~ Sentiment + Politics +
    NewsCountry + USCompany + CompanyIndustry +
    Total_Assets_log + Total_Employees_log,
  data = NewsEfuelCompanyEfuel
)
summary(lm9)

vif(lm9, type = "predictor")
alias(lm9)


#EV: No Newspaper, DisasterType, DisasterSubgroup, CompanyIndustry
```

```
plm10_sent <- plm(
  CAR_winsor ~ Sentiment +
    NewsCountry + USCompany +
    Total_Assets_log + Total_Employees_log + Year,
  data = NewsEVCompanyEV,
  index = c("Company", "EventDate", "Headline"),
  model = "within",          # "within" = fixed effects
  effect = "individual"      # or "twoways" if you want time FE too
)
summary(plm10_sent)

plm10_pol <- plm(
  CAR_winsor ~ Politics +
    NewsCountry + USCompany +
    Total_Assets_log + Total_Employees_log + Year,
  data = NewsEVCompanyEV,
  index = c("Company", "EventDate", "Headline"),
  model = "within",          # "within" = fixed effects
  effect = "individual"      # or "twoways" if you want time FE too
)
summary(plm10_pol)

plm10 <- plm(
  CAR_winsor ~ Sentiment + Politics + Sentiment * Politics +
    NewsCountry + USCompany +
    Total_Assets_log + Total_Employees_log + Year,
  data = NewsEVCompanyEV,
  index = c("Company", "EventDate", "Headline"),
  model = "within",          # "within" = fixed effects
  effect = "individual"      # or "twoways" if you want time FE too
)
summary(plm10)

#CompanyIndustry only has 1 level: EV manufacturers, therefore remove that variable
sapply(NewsEVCompanyEV[, c("Sentiment", "Politics", "NewsCountry",
                           "USCompany", "CompanyIndustry", "Year")],
       function(x) length(unique(x)))

lm10 <- lm(
  CAR_winsor ~ Sentiment + Politics + Sentiment * Politics +
    NewsCountry + USCompany +
    Total_Assets_log + Total_Employees_log + Year,
  data = NewsEVCompanyEV
)
summary(lm10)

vif(lm10, type = "predictor")
alias(lm10)

stargazer(plm6, plm7, plm8, plm9, plm10,
          type = "latex",   # use "latex" for LaTeX output and "text" for text output
          title = "Panel Regression Results Specific Technologies",
          column.labels = c("Hydrogen" , "Ammonia", "Methanol", "E-fuel", "EV"),
          dep.var.labels = "CAR_0_10_winsor",
          model.names = FALSE,
          omit.stat = c("f", "ser"),
          no.space = TRUE,
          digits = 3)
```

```
###Fully integrated model
# Create a copy for the matched subset
df_matched <- df

# Assign technology group based on matched news and company tech
df_matched$TechMatchGroup[df_matched$NewsHydrogen == 1 & df_matched$CompanyHydrogen == 1] <-
"Hydrogen"
df_matched$TechMatchGroup[df_matched$NewsAmmonia == 1 & df_matched$CompanyAmmonia == 1] <-
"Ammonia"
df_matched$TechMatchGroup[df_matched$NewsMethanol == 1 & df_matched$CompanyMethanol == 1] <-
"Methanol"
df_matched$TechMatchGroup[df_matched$NewsEfuel == 1 & df_matched$CompanyEfuel == 1] <- "Efuel"
df_matched$TechMatchGroup[df_matched$NewsEV == 1 & df_matched$CompanyEV == 1] <- "EV"

# Keep only rows where a match was found
df_matched <- df_matched %>% filter(!is.na(TechMatchGroup))

# Convert to factor
df_matched$TechMatchGroup <- factor(df_matched$TechMatchGroup)

plm_tech_interacted <- plm(
  CAR_winsor ~ (Sentiment + Politics) * TechMatchGroup +
    NewsCountry + USCompany +
    Total_Assets_log + Total_Employees_log + Year,
  data = df_matched,
  index = c("Company", "EventDate", "Headline"),
  model = "within",
  effect = "individual"
)
summary(plm_tech_interacted)

lm_tech_interacted <- lm(
  CAR_winsor ~ (Sentiment + Politics) * TechMatchGroup +
    NewsCountry + USCompany +
    Total_Assets_log + Total_Employees_log + Year,
  data = df_matched,
)
summary(lm_tech_interacted)

vif(lm_tech_interacted, type = "predictor")
alias(lm_tech_interacted)

stargazer(plm_tech_interacted,
          type = "latex",   # use "latex" for LaTeX output and "text" for text output
          title = "Panel Regression Results Integrated Model",
          dep.var.labels = "CAR_0_10_winsor",
          model.names = FALSE,
          omit.stat = c("f", "ser"),
          no.space = TRUE,
          digits = 3)

#______________________________________________________________________________

###Summary Statistics

#______________________________________________________________________________

###Part 1: CAR_0_10 normal & winsor
summary(df$CAR_0_10)
sd(df$CAR_0_10, na.rm = TRUE)
```

```
summary(df$CAR_winsor)
sd(df$CAR_winsor, na.rm = TRUE)

###Part 2: Regression + control variables
#Politics & Sentiment
desired_order <- c("Left", "Center", "Right", "Far Right")
counts_df <- table(df$Sentiment, df$Politics)[, desired_order]
proportions_df <- prop.table(counts_df, margin = 2)

# Combine into formatted strings: "count (proportion)"
formatted_matrix_df <- matrix(
  paste0(counts_df, " (", sprintf("%.2f", proportions_df), ")"),
  nrow = nrow(counts_df),
  dimnames = list(rownames(counts_df), colnames(counts_df))
)

# Add row totals
row_totals_df <- rowSums(counts_df)
row_total_col_df <- paste0(row_totals_df, " (", sprintf("%.2f", row_totals_df / sum(row_totals_df)), ")")
formatted_matrix_with_row_total_df <- cbind(formatted_matrix_df, Total = row_total_col_df)

# Add column totals
col_totals_df <- colSums(counts_df)
col_total_row_df <- paste0(col_totals_df, " (", sprintf("%.2f", col_totals_df / sum(col_totals_df)), ")")
total_total_df <- paste0(sum(counts_df), " (1.00)")
formatted_matrix_with_totals_df <- rbind(formatted_matrix_with_row_total_df, Total = c(col_total_row_df,
total_total_df))

# View result
print(formatted_matrix_with_totals_df)

#Company Control variables
summary(df$Total_Assets)      #check if log-transform needed
sd(df$Total_Assets, na.rm = TRUE)

summary(df$Total_Employees)   #check if log-transform needed
sd(df$Total_Employees, na.rm = TRUE)

summary(df$Total_Assets_log)
sd(df$Total_Assets_log, na.rm = TRUE)

summary(df$Total_Employees_log)
sd(df$Total_Employees_log, na.rm = TRUE)

table(df$USCompany)
prop.table(table(df$USCompany))

table(df$CompanyIndustry)
prop.table(table(df$CompanyIndustry))

#News control variables
table(df$NewsCountry)
prop.table(table(df$NewsCountry))

df %>%
  count(Newspaper, sort = TRUE) %>%   #Newspaper counts: top 5
  top_n(5)
table(df$Newspaper)
prop.table(table(df$Newspaper))
```

```
df %>%
  count(MonthYear) %>%       # Count of events per year-month
  arrange(desc(n)) %>%
  head(10)

#Time control variables
df %>%
  count(Year) %>%       # Count of events per year-month
  arrange(desc(n)) %>%
  head(10)

df %>%
  count(Year) %>%
  mutate(Proportion = n / sum(n)) %>%
  arrange(desc(n)) %>%
  head(10)

#Disaster control variables
table(df$DisasterType)
table(df$DisasterSubgroup)


###Part 3: Technology Specific
binary_cols <- c("NewsAmmonia", "NewsMethanol", "NewsHydrogen", "NewsEfuel", "NewsEV",
                 "CompanyAmmonia", "CompanyMethanol", "CompanyHydrogen", "CompanyEfuel", "CompanyEV",
                 "MatchAmmonia", "MatchMethanol", "MatchHydrogen", "MatchEfuel", "MatchEV")

# Compute counts and proportions
binary_summary <- lapply(binary_cols, function(col) {
  counts <- table(df[[col]])
  proportions <- prop.table(counts)
  data.frame(
    Variable = col,
    Value = names(counts),
    Count = as.integer(counts),
    Proportion = round(as.numeric(proportions), 4)
  )
})

# Combine into a single data frame
binary_summary_df <- do.call(rbind, binary_summary)

# Print the result
print(binary_summary_df)
```

Appendix 4: Python Code

```
import kaggle

### Initialize the Kaggle API
api = kaggle.KaggleApi()
api.authenticate()

# Specify your dataset path and target download folder
dataset_path = "jordankrishnayah/45m-headlines-from-2007-2022-10-largest-sites"
target_path = "c:\\Users\\clara\\OneDrive - UCL\\CEMS\\Thesis\\Thesis Python_Code"

### Download the dataset
api.dataset_download_files(dataset_path, path=target_path, unzip=True)

print("Dataset downloaded to:", target_path)

### import dataset into python
import pandas as pd

# Example: Assuming one of the files is named 'data.csv'
file_path = target_path + '\\headlines.csv'
data = pd.read_csv(file_path)

# Display the first few rows of the dataframe
print(data.head())

# Display basic statistics about the data
print(data.describe())

###1. filter on keywords around clean tech
###keywords = ['clean tech', 'clean technology' , 'green technology', 'renewable energy', 'sustainable energy',
'green power', 'power-to-x', 'solar power', 'solar panel', 'wind power', 'wind turbine', 'wind farm', 'wind mill',
'hydropower', 'bioenergy', 'biofuel', 'geothermal power', 'electric vehicle', 'EV charging', 'EV', 'battery storage',
'energy storage', 'lithium battery', 'smart grid', 'power-to-gas', 'power-to-liquid', 'hydrogen', 'methanol', 'ammonia',
'hydroelectric', 'alternative fuels', 'fuel cells']

keywords = [
    r'\bclean tech\b', r'\bclean technology\b', r'\bgreen technology\b',
    r'\brenewable energy\b', r'\bsustainable energy\b', r'\bgreen power\b',
    r'\bpower-to-x\b', r'\bsolar power\b', r'\bsolar panel\b',
    r'\bwind power\b', r'\bwind turbine\b', r'\bwind farm\b', r'\bwind mill\b',
    r'\bhydropower\b', r'\bbioenergy\b', r'\bbiofuel\b',
    r'\bgeothermal power\b', r'\belectric vehicle\b', r'\bEV charging\b',
    r'\bEV\b', r'\bbattery storage\b', r'\benergy storage\b',
    r'\blithium battery\b', r'\bsmart grid\b', r'\bpower-to-gas\b',
    r'\bpower-to-liquid\b', r'\bhydrogen\b', r'\bmethanol\b', r'\bammonia\b',
    r'\bhydroelectric\b', r'\balternative fuels\b', r'\bfuel cells\b',
    r'\bmicrogrid\b', r'\bcarbon capture\b', r'\benergy efficiency\b',
    r'\blow-carbon technology\b', r'\bsustainable transport\b',
    r'\brenewable resources\b', r'\bwaste-to-energy\b',
    r'\bwave energy\b', r'\bbiogas\b',
    r'\bperovskite solar cells\b',
    r'\bnet zero emissions\b', r'\bsmart energy systems\b',
    r'\bsustainable development\b', r'\benergy transition\b',
    r'\bgreen building\b', r'\bsmart thermostats\b',
    r'\benergy management systems\b'
]

# Create a filter that checks if any of the keywords are in the 'Headline' column
```

```
# The str.contains method can take a regex pattern, so join the keywords into a regex pattern that matches any of
them
# The na=False parameter makes sure that NaN values return False
filtered_data = data[data['Headline'].str.contains('|'.join(keywords), case=False, na=False)]

# Define the path to save the filtered dataset
output_path = r"c:\\Users\\clara\\OneDrive - UCL\\CEMS\\Thesis\\Thesis Python_Code\\filtered_headlines.csv"

# Save the filtered data to a new CSV file
filtered_data.to_csv(output_path, index=False)

print(f"Filtered data saved to: {output_path}")

# Remove duplicate rows
# You can adjust subset=['Headline'] to consider other columns or remove it to consider all columns
unique_data = filtered_data.drop_duplicates(subset=['Headline'], keep='first')

# Define the path to save the filtered dataset
output_path = r"c:\\Users\\clara\\OneDrive - UCL\\CEMS\\Thesis\\Thesis
Python_Code\\filtered_unique_headlines.csv"

# Save the filtered data to a new CSV file
unique_data.to_csv(output_path, index=False)

print(f"Unique filtered data saved to: {output_path}")


### 2. Sentiment analysis using textblok nltk
import nltk
nltk.download('vader_lexicon')
from textblob import TextBlob
import pandas as pd

from nltk.sentiment import SentimentIntensityAnalyzer

# Define the path to your dataset
file_path = r"c:\\Users\\clara\\OneDrive - UCL\\CEMS\\Thesis\\Thesis
Python_Code\\filtered_unique_headlines.csv"

# Load the dataset
data = pd.read_csv(file_path)

# Initialize VADER
sia = SentimentIntensityAnalyzer()

# Function to get sentiment category
def get_sentiment_category(text):
    scores = sia.polarity_scores(text)
    if scores['compound'] >= 0.05:
        return 'Positive'
    elif scores['compound'] <= -0.05:
        return 'Negative'
    else:
        return 'Neutral'

# Apply the sentiment analysis
data['Sentiment Category'] = data['Headline'].apply(get_sentiment_category)

# Define the path to save the sentiment analysis results
output_path = r"c:\Users\clara\OneDrive - UCL\CEMS\Thesis\Thesis Python_Code\sentiment_headlines.csv"
```

```
# Save the results to a new CSV file
data.to_csv(output_path, index=False)

print(f"Sentiment analysis completed and saved to: {output_path}")

### Sentiment analysis based on entire text
import pandas as pd
import requests
from bs4 import BeautifulSoup
from nltk.sentiment import SentimentIntensityAnalyzer

# Load dataset with URLs
file_path = r"c:\\Users\\clara\\OneDrive - UCL\\CEMS\\Thesis\\Thesis
Python_Code\\filtered_unique_headlines.csv"
data = pd.read_csv(file_path)

# Initialize VADER
sia = SentimentIntensityAnalyzer()

# Function to scrape article text
def get_article_text(url):
    try:
        response = requests.get(url)
        soup = BeautifulSoup(response.content, 'html.parser')
        article_text = ' '.join(p.get_text() for p in soup.find_all('p'))
        return article_text
    except Exception as e:
        return None

# Function to get sentiment category
def get_sentiment_category(text):
    scores = sia.polarity_scores(text)
    return 'Positive' if scores['compound'] >= 0.05 else 'Negative' if scores['compound'] <= -0.05 else 'Neutral'

# Apply text scraping and sentiment analysis
data['Article Text'] = data['URL'].apply(get_article_text)
data['Sentiment Category'] = data['Article Text'].dropna().apply(get_sentiment_category)

# Save the results
output_path = r"c:\Users\clara\OneDrive - UCL\\CEMS\\Thesis\\Thesis
Python_Code\sentiment_full_articles.csv"
data.to_csv(output_path, index=False)

print(f"Sentiment analysis of full articles completed and saved to: {output_path}")
```

Appendix 5: CompanyData Description and Source

Verdonck, C. (2025e). *CompanyDataFINAL* [Unpublished raw data].

The sample list of 108 companies compiled via S&P Capital IQ (S&P Global Market Intelligence, n.d).

```
  Company       CompanyName  CompanyCountry USCompany CompanyIndustryBroad CompanyIndustry CompanyDescription ProductDescription CompanyEV
  <chr>         <chr>        <chr>              <dbl> <chr>                <chr>           <chr>              <chr>                  <dbl>
1 NYSE:AYI      Acuity Bran… U.S.                   1 Electrical Equipmen… Energy Technol… Acuity Brands, In… "Building Managem…          0
2 NasdaqCM:ADN  Advent Tech… U.S.                   1 Electrical Equipmen… E-fuel          Advent Technologi… "Electrodes:\r\nE…          0
3 AIM:AFC       AFC Energy … UK                     0 Electrical Equipmen… E-fuel          AFC Energy plc en… "Electric Vehicle…          0
4 NYSE:APD      Air Product… U.S.                   1 Chemicals (Primary)  Energy Technol… Air Products and … "Acetylenic Alcoh…          0
5 LSE:ALK       Alkemy Capi… UK                     0 Electrical Equipmen… Energy Storage  Alkemy Capital In… "Construction Ser…          0
6 OB:ALT        Alternus En… U.S.                   1 Electric Power By S… Solar Energy    Alternus Energy G… "Electricity Gene…          0
# ℹ 4 more variables: CompanyAmmonia <dbl>, CompanyMethanol <dbl>, CompanyHydrogen <dbl>, CompanyEfuel <dbl>
```

Appendix 6: StockData Description and Source

Verdonck, C. (2025b). *StockDataFINAL2025* [Unpublished raw data].

The stock market closing prices of the sample list of companies of the CompanyDataset, from the Bloomberg Terminal (2006-2022) (Bloomberg L.P. (n.d.).

```
  Date                `NYSE:AYI` `NasdaqCM:ADN` `AIM:AFC` `LSE:ALK` `OB:ALT` `NYSE:AMPS` `NasdaqGS:AMSC` `NYSE:AME` `NYSE:AMPX`
  <dttm>                   <dbl>          <dbl>     <dbl>     <dbl>    <dbl>       <dbl>           <dbl>      <dbl>       <dbl>
1 2006-01-03 00:00:00       26.6             NA        NA        NA    2100.          NA            84.4       12.6          NA
2 2006-01-04 00:00:00       27.1             NA        NA        NA    2115.          NA            89.2       12.7          NA
3 2006-01-05 00:00:00       29.1             NA        NA        NA    2115.          NA            88.4       13.0          NA
4 2006-01-06 00:00:00       29.3             NA        NA        NA    2088.          NA            88.4       13.2          NA
5 2006-01-09 00:00:00       30.2             NA        NA        NA    2097.          NA            92.5       13.5          NA
6 2006-01-10 00:00:00       30.4             NA        NA        NA    2115.          NA            92.3       13.3          NA
# ℹ 209 more variables: `NasdaqGM:ARRY` <dbl>, `LSE:AML` <dbl>, `NYSE:ATKR` <dbl>, `NasdaqCM:AYRO` <dbl>, `NYSE:BW` <dbl>,
#   `NasdaqCM:BEEM` <dbl>, `NasdaqCM:BLNK` <dbl>, `NYSE:BE` <dbl>, `NasdaqCM:BWEN` <dbl>, `LSE:CWR` <dbl>, `NYSE:CHPT` <dbl>,
#   `NasdaqGM:CSLR` <dbl>, `NasdaqGM:CNTM` <dbl>, `LSE:DIA` <dbl>, `LSE:DSCV` <dbl>, `NasdaqCM:DFLI` <dbl>, `LSE:DRX` <dbl>,
#   `NasdaqCM:WATT` <dbl>, `NasdaqCM:EFOI` <dbl>, `NYSE:NRGV` <dbl>, `NYSE:ENS` <dbl>, `NasdaqGS:ENVX` <dbl>, `NasdaqCM:EVTV` <dbl>,
#   `NasdaqCM:EOSE` <dbl>, `NYSE:GWH` <dbl>, `NasdaqCM:XPON` <dbl>, `NasdaqCM:FFIE` <dbl>, `NasdaqGS:FLNC` <dbl>, `NasdaqCM:FLUX` <dbl>,
#   `NasdaqCM:FTCI` <dbl>, `NasdaqGM:FCEL` <dbl>, `AIM:GELN` <dbl>, `NYSE:GNK` <dbl>, `NYSE:GNRC` <dbl>, `NYSE:GM` <dbl>,
#   `NYSE:HUBB` <dbl>, `NasdaqCM:IPWR` <dbl>, `AIM:IKA` <dbl>, `AIM:IES` <dbl>, `AIM:ITM` <dbl>, `NYSE:KEX` <dbl>, …
```

Appendix 7: SPXData Description and Source

Verdonck, C. (2025c). *SPX* [Unpublished raw data].

S&P500 market price data, from the Bloomberg Terminal (2006-2022) (Bloomberg L.P., n.d.).

```
  Date                SPXIndex SPX_Return_notlog
  <dttm>                 <dbl>             <dbl>
1 2006-01-04 00:00:00    1273.        NA        
2 2006-01-05 00:00:00    1273.         0.0000157
3 2006-01-06 00:00:00    1285.         0.00940  
4 2006-01-09 00:00:00    1290.         0.00366  
5 2006-01-10 00:00:00    1290.        -0.000357 
6 2006-01-11 00:00:00    1294.         0.00348  
```

Appendix 8: RFRData Description and Source

Verdonck, C. (2025d). *RFR* [Unpublished raw data].

The risk-free-rate data, from the Federal Reserve Bank of St. Louis (2006-2022) (Federal Reserve Bank of St. Louis., n.d.).

```
  Date                   RFR RFR_daily
  <dttm>               <dbl>     <dbl>
1 2006-01-03 00:00:00 0.0437  0.000170
2 2006-01-04 00:00:00 0.0436  0.000169
3 2006-01-05 00:00:00 0.0436  0.000169
4 2006-01-06 00:00:00 0.0438  0.000170
5 2006-01-09 00:00:00 0.0438  0.000170
6 2006-01-10 00:00:00 0.0443  0.000172
```

Appendix 9: NewspaperData Description and Source

Verdonck, C. (2025a). *NewspaperData* [Unpublished raw data].

The newspaper data of U.S. and UK news headlines, from Kaggle based on most popular newspapers on Statista (2007-2022) (Kaggle, 2023; Statista, 2025e).

```
  EventDate           Newspaper      NewsCountry Headline         Sentiment Politics NewsEV NewsAmmonia NewsMethanol NewsHydrogen NewsEfuel
  <dttm>              <chr>                <dbl> <chr>            <chr>     <chr>     <dbl>       <dbl>        <dbl>        <dbl>     <dbl>
1 2007-01-24 00:00:00 New York Times           1 Bush Seeks Va…   Positive  Left          0           0            0            0         1
2 2007-04-28 00:00:00 New York Times           1 Hydrogen's Se…   Positive  Left          0           0            0            1         0
3 2007-04-28 00:00:00 New York Times           1 Hydrogen's Se…   Negative  Left          0           0            0            1         0
4 2007-04-28 00:00:00 New York Times           1 Hydrogen's Se…   Positive  Left          0           0            0            1         0
5 2007-05-01 00:00:00 New York Times           1 Hydrogen's Se…   Neutral   Left          0           0            0            1         0
6 2007-05-12 00:00:00 CNBC                     1 Hydrogen Cars…   Neutral   Center        0           0            0            1         0
```

Appendix 10: Control Variables Data Description and Source

Verdonck, C. (2025f). *BS_TOT_ASSETSCompanies* [Unpublished raw data].

BS_TOT_ASSETS, from the Bloomberg Terminal (2006-2022) (Bloomberg L.P., n.d.).

```
  Date                 Year `NYSE:AYI` `NasdaqCM:ADN` `AIM:AFC` `LSE:ALK` `OB:ALT` `NYSE:AMPS` `NasdaqGS:AMSC` `NYSE:AME` `NYSE:AMPX`
  <dttm>              <dbl>      <dbl>          <dbl>     <dbl>     <dbl>    <dbl>       <dbl>           <dbl>      <dbl>       <dbl>
1 2006-08-31 00:00:00  2006      1444.             NA    NA            NA     NA            NA            133.      2131.          NA
2 2007-08-31 00:00:00  2007      1618.             NA      3.36        NA     66.6          NA            132.      2746.          NA
3 2008-08-31 00:00:00  2008      1409.             NA      5.14        NA     56.4          NA            261.      3056.          NA
4 2009-08-31 00:00:00  2009      1291.             NA      3.00        NA     34.1          NA            309.      3246.          NA
5 2010-08-31 00:00:00  2010      1504.             NA      6.99        NA     27.2          NA            400.      3819.          NA
6 2011-08-31 00:00:00  2011      1597.             NA      7.78        NA     22.8          NA            441.      4319.          NA
# ℹ 100 more variables: `NasdaqGM:ARRY` <dbl>, `LSE:AML` <dbl>, `NYSE:ATKR` <dbl>, `NasdaqCM:AYRO` <dbl>, `NYSE:BW` <dbl>,
#   `NasdaqCM:BEEM` <dbl>, `NasdaqCM:BLNK` <dbl>, `NYSE:BE` <dbl>, `NasdaqCM:BWEN` <dbl>, `LSE:CWR` <dbl>, `NYSE:CHPT` <dbl>,
#   `NasdaqGM:CSLR` <dbl>, `NasdaqGM:CNTM` <dbl>, `LSE:DIA` <dbl>, `LSE:DSCV` <dbl>, `NasdaqCM:DFLI` <dbl>, `LSE:DRX` <dbl>,
#   `NasdaqCM:WATT` <dbl>, `NasdaqCM:EFOI` <dbl>, `NYSE:NRGV` <dbl>, `NYSE:ENS` <dbl>, `NasdaqGS:ENVX` <dbl>, `NasdaqCM:EVTV` <dbl>,
#   `NasdaqCM:EOSE` <dbl>, `NYSE:GWH` <dbl>, `NasdaqCM:XPON` <dbl>, `NasdaqCM:FFIE` <dbl>, `NasdaqGS:FLNC` <dbl>, `NasdaqCM:FLUX` <dbl>,
#   `NasdaqCM:FTCI` <dbl>, `NasdaqGM:FCEL` <dbl>, `AIM:GELN` <dbl>, `NYSE:GNK` <dbl>, `NYSE:GNRC` <dbl>, `NYSE:GM` <dbl>,
#   `NYSE:HUBB` <dbl>, `NasdaqCM:IPWR` <dbl>, `AIM:IKA` <dbl>, `AIM:IES` <dbl>, `AIM:ITM` <dbl>, `NYSE:KEX` <dbl>, …
```

Verdonck, C. (2025g). *NUM_OF_EMPLOYEESCompanies* [Unpublished raw data].

NUM_OF_EMPLOYEES, from the Bloomberg Terminal (2006-2022) (Bloomberg L.P., n.d.).

```
  Date                 Year `NYSE:AYI` `NasdaqCM:ADN` `AIM:AFC` `LSE:ALK` `OB:ALT` `NYSE:AMPS` `NasdaqGS:AMSC` `NYSE:AME` `NYSE:AMPX`
  <dttm>              <dbl>      <dbl>          <dbl>     <dbl>     <dbl>    <dbl>       <dbl>           <dbl>      <dbl>       <dbl>
1 2006-08-31 00:00:00  2006      10600             NA        NA        NA        3          NA             241      10400          NA
2 2007-08-31 00:00:00  2007      10000             NA        16        NA       91          NA             263      11300          NA
3 2008-08-31 00:00:00  2008       6500             NA        26        NA      151          NA             382      11700          NA
4 2009-08-31 00:00:00  2009       6000             NA        24        NA      147          NA             519      10100          NA
5 2010-08-31 00:00:00  2010       6000             NA        25        NA       85          NA             714      11600          NA
6 2011-08-31 00:00:00  2011       6000             NA        31        NA       73          NA             848      12200          NA
# ℹ 100 more variables: `NasdaqGM:ARRY` <dbl>, `LSE:AML` <dbl>, `NYSE:ATKR` <dbl>, `NasdaqCM:AYRO` <dbl>, `NYSE:BW` <dbl>,
#   `NasdaqCM:BEEM` <dbl>, `NasdaqCM:BLNK` <dbl>, `NYSE:BE` <dbl>, `NasdaqCM:BWEN` <dbl>, `LSE:CWR` <dbl>, `NYSE:CHPT` <dbl>,
#   `NasdaqGM:CSLR` <dbl>, `NasdaqGM:CNTM` <dbl>, `LSE:DIA` <dbl>, `LSE:DSCV` <dbl>, `NasdaqCM:DFLI` <dbl>, `LSE:DRX` <dbl>,
#   `NasdaqCM:WATT` <dbl>, `NasdaqCM:EFOI` <dbl>, `NYSE:NRGV` <dbl>, `NYSE:ENS` <dbl>, `NasdaqGS:ENVX` <dbl>, `NasdaqCM:EVTV` <dbl>,
#   `NasdaqCM:EOSE` <dbl>, `NYSE:GWH` <dbl>, `NasdaqCM:XPON` <dbl>, `NasdaqCM:FFIE` <dbl>, `NasdaqGS:FLNC` <dbl>, `NasdaqCM:FLUX` <dbl>,
#   `NasdaqCM:FTCI` <dbl>, `NasdaqGM:FCEL` <dbl>, `AIM:GELN` <dbl>, `NYSE:GNK` <dbl>, `NYSE:GNRC` <dbl>, `NYSE:GM` <dbl>,
#   `NYSE:HUBB` <dbl>, `NasdaqCM:IPWR` <dbl>, `AIM:IKA` <dbl>, `AIM:IES` <dbl>, `AIM:ITM` <dbl>, `NYSE:KEX` <dbl>, …
```

Appendix 11: Complete Variable Description Used in Regression Analyses

Verdonck, C. (2025a). *NewspaperData* [Unpublished raw data].

Verdonck, C. (2025b). *StockDataFINAL2025* [Unpublished raw data].

Verdonck, C. (2025c). *SPX* [Unpublished raw data].

Verdonck, C. (2025d). *RFR* [Unpublished raw data].

Verdonck, C. (2025e). *CompanyDataFINAL* [Unpublished raw data].

Verdonck, C. (2025f). *BS_TOT_ASSETSCompanies* [Unpublished raw data].

Verdonck, C. (2025g). *NUM_OF_EMPLOYEESCompanies* [Unpublished raw data].

Verdonck, C. (2025h). *EMDATData* [Unpublished raw data].

```
 $ EventDate          : Factor w/ 1034 levels "2007-01-24","2007-04-28",..: 1 1 1 1 1 1 1 1 1 1 ...
 $ Newspaper          : Factor w/ 10 levels "BBC","CNBC","CNN",..: 7 7 7 7 7 7 7 7 7 7 ...
 $ NewsCountry        : Factor w/ 2 levels "0","1": 2 2 2 2 2 2 2 2 2 2 ...
 $ Headline           : chr [1:128372] "Bush Seeks Vast, Mandatory Increase in Alternative Fuels and Greater Vehicle Efficiency" "Bush Seeks
Vast, Mandatory Increase in Alternative Fuels and Greater Vehicle Efficiency" "Bush Seeks Vast, Mandatory Increase in Alternative Fuels and G
reater Vehicle Efficiency" "Bush Seeks Vast, Mandatory Increase in Alternative Fuels and Greater Vehicle Efficiency" ...
 $ Sentiment          : Factor w/ 3 levels "Neutral","Negative",..: 3 3 3 3 3 3 3 3 3 3 ...
 $ Politics           : Factor w/ 4 levels "Center","Far Right",..: 3 3 3 3 3 3 3 3 3 3 ...
 $ NewsEV             : num [1:128372] 0 0 0 0 0 0 0 0 0 0 ...
 $ NewsAmmonia        : num [1:128372] 0 0 0 0 0 0 0 0 0 0 ...
 $ NewsMethanol       : num [1:128372] 0 0 0 0 0 0 0 0 0 0 ...
 $ NewsHydrogen       : num [1:128372] 0 0 0 0 0 0 0 0 0 0 ...
 $ NewsEfuel          : num [1:128372] 1 1 1 1 1 1 1 1 1 1 ...
 $ CompanyID          : int [1:128372] 3 5 6 8 11 12 13 14 16 17 ...
 $ Company            : Factor w/ 108 levels "AIM:AFC","AIM:GELN",..: 3 5 6 8 11 12 13 14 64 65 ...
 $ CAR_0_10           : num [1:128372] 0.01646 0.27769 -0.07428 0.01252 0.00825 ...
 $ CAR_winsor         : num [1:128372] 0.01646 0.27769 -0.07428 0.01252 0.00825 ...
 $ Year               : Factor w/ 19 levels "2007","2008",..: 1 1 1 1 1 1 1 1 1 1 ...
 $ MonthYear          : Factor w/ 180 levels "2007-01","2007-04",..: 1 1 1 1 1 1 1 1 1 1 ...
 $ Total_Assets       : num [1:128372] 69.492 33.012 0.248 212.08 13.665 ...
 $ Total_Employees    : num [1:128372] 147 37 NA 9814 52 ...
 $ CompanyCountry     : chr [1:128372] "UK" "UK" "UK" "UK" ...
 $ CompanyName        : chr [1:128372] "Invinity Energy Systems plc" "ITM Power Plc" "PowerHouse Energy Group Plc" "Volex plc" ...
 $ USCompany          : Factor w/ 2 levels "0","1": 1 1 1 1 1 2 1 1 2 2 ...
 $ CompanyIndustry    : Factor w/ 8 levels "E-fuel","Energy Storage",..: 2 1 3 3 1 3 6 2 3 3 ...
 $ CompanyEV          : num [1:128372] 0 0 0 0 0 0 0 0 0 0 ...
 $ CompanyAmmonia     : num [1:128372] 0 1 0 0 1 0 0 0 0 0 ...
 $ CompanyMethanol    : num [1:128372] 0 1 0 0 1 0 0 0 0 0 ...
 $ CompanyHydrogen    : num [1:128372] 0 1 1 0 1 0 0 0 1 0 ...
 $ CompanyEfuel       : num [1:128372] 0 1 1 0 1 0 0 0 0 0 ...
 $ EndDate            : POSIXct[1:128372], format: NA NA NA NA ...
 $ DisasterLength     : num [1:128372] NA NA NA NA NA NA NA NA NA NA ...
 $ DisasterGroup      : chr [1:128372] NA NA NA NA ...
 $ DisasterSubgroup   : Factor w/ 4 levels "Climatological",..: NA NA NA NA NA NA NA NA NA NA ...
 $ DisasterType       : Factor w/ 6 levels "Drought","Earthquake",..: NA NA NA NA NA NA NA NA NA NA ...
 $ DisasterSubtype    : chr [1:128372] NA NA NA NA ...
 $ EventName          : chr [1:128372] NA NA NA NA ...
 $ ISO                : chr [1:128372] NA NA NA NA ...
 $ Subregion          : chr [1:128372] NA NA NA NA ...
 $ Location           : chr [1:128372] NA NA NA NA ...
 $ Origin             : chr [1:128372] NA NA NA NA ...
 $ StartYear          : Factor w/ 14 levels "2007","2008",..: NA NA NA NA NA NA NA NA NA NA ...
 $ StartMonth         : num [1:128372] NA NA NA NA NA NA NA NA NA NA ...
 $ StartDay           : num [1:128372] NA NA NA NA NA NA NA NA NA NA ...
 $ EndYear            : num [1:128372] NA NA NA NA NA NA NA NA NA NA ...
 $ EndMonth           : num [1:128372] NA NA NA NA NA NA NA NA NA NA ...
 $ EndDay             : num [1:128372] NA NA NA NA NA NA NA NA NA NA ...
 $ Total_Assets_log   : num [1:128372] 4.255 3.527 0.222 5.362 2.685 ...
 $ Total_Employees_log: num [1:128372] 4.99 3.61 NA 9.19 3.95 ...
 $ MatchEV            : num [1:128372] 0 0 0 0 0 0 0 0 0 0 ...
 $ MatchAmmonia       : num [1:128372] 0 0 0 0 0 0 0 0 0 0 ...
 $ MatchMethanol      : num [1:128372] 0 0 0 0 0 0 0 0 0 0 ...
 $ MatchHydrogen      : num [1:128372] 0 0 0 0 0 0 0 0 0 0 ...
 $ MatchEfuel         : num [1:128372] 0 1 1 0 1 0 0 0 0 0 ...
 $ NewsNoAmmonia      : num [1:128372] 1 1 1 1 1 1 1 1 1 1 ...
 $ NewsNoMethanol     : num [1:128372] 1 1 1 1 1 1 1 1 1 1 ...
 $ NewsNoHydrogen     : num [1:128372] 1 1 1 1 1 1 1 1 1 1 ...
 $ NewsNoEfuel        : num [1:128372] 0 0 0 0 0 0 0 0 0 0 ...
 $ NewsNoEV           : num [1:128372] 1 1 1 1 1 1 1 1 1 1 ...
 - attr(*, "na.action")= 'omit' Named int [1:41450] 1 2 4 7 9 10 15 18 19 21 ...
  ..- attr(*, "names")= chr [1:41450] "1" "2" "4" "7" ...
```