

Media Attention to Environmental Issues and ESG Investing*

Balázs J. Csillag – Marcell P. Granát – Gábor Neszveda

We analyse how ESG scores affect future returns when environmental issues receive higher media coverage. Investors might take environmental aspects into account if they are confronted with the issue of global warming more frequently in the press. We assess the prevalence of environmental issues in the media with a machine learning-based Structural Topic Modelling (STM) methodology, using a news archive published in the USA. Running Fama-MacBeth regressions, we find that in periods when the media actively report on environmental issues, ESG scores have a significant negative impact on future returns, whereas, in months when fewer such articles are published, investors do not take sustainability measures into account, and ESG scores have no explanatory power.

Journal of Economic Literature (JEL) codes: C55, G12

Keywords: ESG, environmental issues, investors' attention, Structural Topic Model, Fama-MacBeth regression

1. Introduction

The study of investments that ensure environmentally responsible and sustainable development is becoming an increasingly relevant research topic in the field of empirical asset pricing. Corporate managers also appear to increasingly recognise the importance of environmental and sustainability issues (Flammer 2013). Environmental, Sustainability and Governance (ESG) scores are a popular and commonly used indicator to measure the commitment of a firm to addressing environmental and social issues (Townsend 2020). ESG scores attempt to measure compliance with these three criteria and offer a proxy for overall company

* The papers in this issue contain the views of the authors which are not necessarily the same as the official views of the Magyar Nemzeti Bank.

Balázs J. Csillag is a Student at John von Neumann University. Email: csillagb3@gmail.com
Marcell P. Granát is an Education and Research Expert at the Magyar Nemzeti Bank, Directorate for International Monetary Policy Analysis and Training of Economic Sciences; and an Assistant Lecturer at John von Neumann University. Email: granatm@mn.b.hu
Gábor Neszveda is a Senior Educational and Research Expert at the Magyar Nemzeti Bank, Directorate for International Monetary Policy Analysis and Training of Economic Sciences; and an Associate Professor at John von Neumann University; Deputy Head of MNB Institute. Email: neszvedag@mn.b.hu

The first version of the English manuscript was received on 13 June 2022.

DOI: <https://doi.org/10.33893/FER.21.4.129>

sustainability. Although numerous competing measures have been proposed, *Talan and Sharma (2019)* and others find that ESG scores are one of the most effective and widely-used indicators.

Being environmentally and socially responsible might be beneficial for a company. Past research provides evidence of a lower cost of capital for companies with higher ESG scores (*Kotsantonis – Serafeim 2019*). Sustainable operations may also lead to higher efficiency (*Gillan et al. 2010*) and positively affect the return on firm equity or the return on assets (*Buallay 2019*). As a result, investors may prefer firms with “sustainable” management, which should be reflected in a positive correlation between ESG scores, stock prices and future returns. On the other hand, if attempts to increase ESG scores distract a company from their primary responsibility to customers and shareholders, this could reduce profits, resulting in a negative correlation between ESG scores and future returns.

Previous research also presents contradictory results. *Sahut and Pasquini-Descomps (2015)* investigate the relationship between ESG scores and stock returns in the US, UK and Swiss markets in the period 2007–2011 and find that ESG scores significantly impact returns negatively only for the United Kingdom. Meanwhile, other sector-specific research (*Buallay 2019*) finds that higher ESG scores among US banks had a significant positive impact on returns between 2007 and 2016. The results suggest there may be country-specific differences in investors’ preferences for ESG investing. Conversely, no significant impact appears at an industry level in the US and Asia-Pacific regions, while investors in Europe appear to be willing to pay a premium for so-called “green stocks” (*Auer – Schuhmacher 2016*). Applying different approaches, several authors (*Jain et al. 2019; Naffa – Fain 2020; Naffa – Fain 2022*) obtain mixed results on the impact of ESG scores on expected returns.

These contradictory results have led to numerous studies on what drives the perceptions and actions of investors with respect to the importance of environmental issues. Due to the limited awareness of people (*Hirshleifer – Teoh 2003; Neszveda 2018*), they are more likely to pay attention to environmental issues if they are confronted with them more frequently or if they experience extreme weather events. Studies (*Li et al. 2011; Akerlof et al. 2013*) report that personal experiences matter a great deal and find that recent experience with global warming (such as extreme weather or natural disasters) increases the perception of climate risk in the United States. *Choi et al. (2020)* determine that retail investors sell carbon-intensive firms in the case of abnormal weather experienced in their surroundings. Recently, extreme weather events appear to trigger more intensive media coverage of environmental issues. *Schmidt (2015)* shows that media attention to climate change increases more in record-breaking warm years than in “normal” years.

In this paper, we analyse how the effect of ESG scores on future returns changes when environmental issues receive relatively high media coverage. People may be more likely to pay attention and react to information that they see more frequently. Various studies have documented that the attention of investors is also limited (*DellaVigna – Pollet 2009; Hirshleifer – Teoh 2003*), and thus they may only take environmental issues into account if they are confronted with the problem of global warming more frequently. Seeing more news on the topic may change their perceptions about the importance of sustainability and cause them to choose or modify their investments accordingly.

Based on these observations, we hypothesise that in periods when the media actively reports on environmental issues ESG scores have a significant impact on future returns. Conversely, in months when fewer such articles are published investors do not take sustainability measures into account, and therefore ESG scores have little or no explanatory power.

We assess the importance of environmental issues in the media applying a *Structural Topic Modelling* (STM) methodology and using a publicly available news archive, which consists of a collection of news reports published on investing.com. The model identifies environmental topics and determines their relative prevalence in news articles.

In our analysis, we run Fama-MacBeth regressions (*Fama – MacBeth 1973*) using observations in months when environmental issues received higher (lower) than average media coverage. These months are referred to as intensive (non-intensive) periods.

However, we find that in periods of intensive environmental media coverage the *Social and Governance Scores in ESG* do significantly affect future returns, while in low-intensity periods they have no explanatory power. These findings are robust to different definitions of intensive and non-intensive periods. Moreover, we find that, using the median as a threshold to define high media-intensity periods, each individual ESG score has significant explanatory power in predicting future returns. These results imply that investors pay more attention to ESG scores when environmental issues receive high media coverage. Conversely, in months when investors are less confronted with environmental problems in the press, they do not take these issues into account to the same extent in their investment decisions.

The study is organised as follows: Section 2 describes the data used in the analysis, data cleaning methods, and provides summary statistics. Section 3 describes the Structural Topic Modelling method and the Fama-MacBeth regression methodology. Section 4 presents our results, while Section 5 summarises our conclusions along with their limitations.

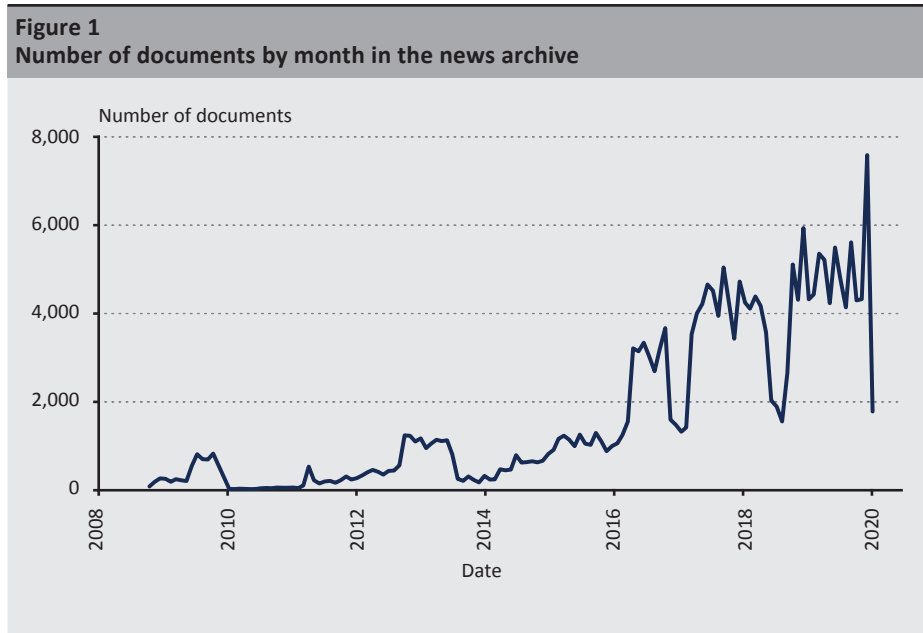
2. Data

In this section, we present in detail the databases used in the analysis. First, we describe stock news data, and then data on company information and stocks' performance. In the last subsection, we provide some key descriptive statistics.

2.1. Stock news data

In addition to the financial data, we need corresponding text data in order to quantify the intensity of the environmental (green) issues in the news at any given time. Therefore, we use a publicly available news archive, which consists of a collection of new reports published on the investing.com website. The data are available at: <https://www.kaggle.com/datasets/gennadiyr/us-equities-news-data>.

Although the dataset contains news from 2008, it is only complete without interruption from 2010 and ends in 2020 (see *Figure 1*). The news is related to US equities which are publicly traded on NYSE/NASDAQ and maintained a price above USD 10 per share through 2020.



In the analysis, we use the daily frequency to generate the prevalence of topics, but then aggregate the intensity of environmental (green) topics present in the articles to a monthly level.

2.2. Stock and company data

We use monthly data on stocks traded on the New York Stock Exchange (NYSE) and the National Association of Securities Dealers Automated Quotations (NASDAQ) stock exchanges. The Refinitiv Datastream (formerly Thomson Reuters Datastream, TDS) was used, instead of the most common Center for Research in Security Prices database (CRSP). *Ince – Porter (2006)* report that after cleaning the inferences drawn from TDS data are similar to those drawn from CRSP.

Our data consist of active and inactive primary equities. To avoid delisting bias (or survival bias) inactive stocks are also included (*Shumway 1997*). One-month US Treasury Bills are used as risk-free returns.¹

The following variables were obtained through TDS:

- *Price*: unadjusted price quoted on the first day of the respective month, data in USD (Datastream database);
- *Total Return Index*: following the literature, we use the total return index as a performance measure, as it adjusts for price movements due to stock splits and dividend payments (Datastream database);
- *Turnover by Volume*: number of stocks traded in the respective month, data in thousands (Datastream database);
- *Common Shares Outstanding*: number of common shares outstanding at the end of the year (Worldscope database);
- *Book Value per Share*: book value per share at the end of the respective fiscal year, data in USD (Worldscope database);
- *Environment Pillar Score (ES)*:² Refinitiv's Environment Pillar Score is the weighted average relative rating of a company based on the reported environmental information and the resulting three environmental category scores;
- *Social Pillar Score (SS)*: Refinitiv's Social Pillar Score is the weighted average relative rating of a company based on the reported social information and the resulting four social category scores;
- *Governance Pillar Score (GS)*: Refinitiv's Governance Pillar Score is the weighted average relative rating of a company based on the reported governance information and the resulting three governance category scores.

¹ Downloaded from Stambaugh's website (<https://finance.wharton.upenn.edu/stambaug>) on 10 March 2022.

² The research uses Refinitiv's own ESG scores as one of the first and most trusted ESG agencies available. However, there may be significant variation in the ESG ratings of different agencies. For this reason, one may obtain different results using other ESG scores.

Using these variables, the following factors were created as these are widely used in asset pricing literature (Carhart 1997):

- *Market beta*: systematic risk. Beta was estimated by running the following OLS regression model on our data. (Using a 36-month rolling window, requiring at least 30 observations per month.)

$$r_{i,t} - r_{f,t} = \alpha + \beta_{i,t}(r_{m,t} - r_{f,t}) + \varepsilon_{i,t} \quad (1)$$

Where $r_{i,t}$ is the return on asset i in period t ; $r_{f,t}$ is the risk-free return in period t ; $r_{m,t}$ is the market return in period t ; α is the intercept of the model; $\beta_{i,t}$ is the market beta of asset i in period t ; and $\varepsilon_{i,t}$ is the error term in period t .

- *Market value (MV)*: the market capitalisation of the stock i in month t

$$MV_{i,t} = \ln(N_{i,t}P_{i,t}) \quad (2)$$

where $N_{i,t}$ is the number of firms i 's common shares outstanding at the end of the year to which month t belongs, $P_{i,t}$ is i 's USD price quoted on the first day of month t .

- *Book-to-market-ratio (BTM)*: stock i 's B/M in month t .

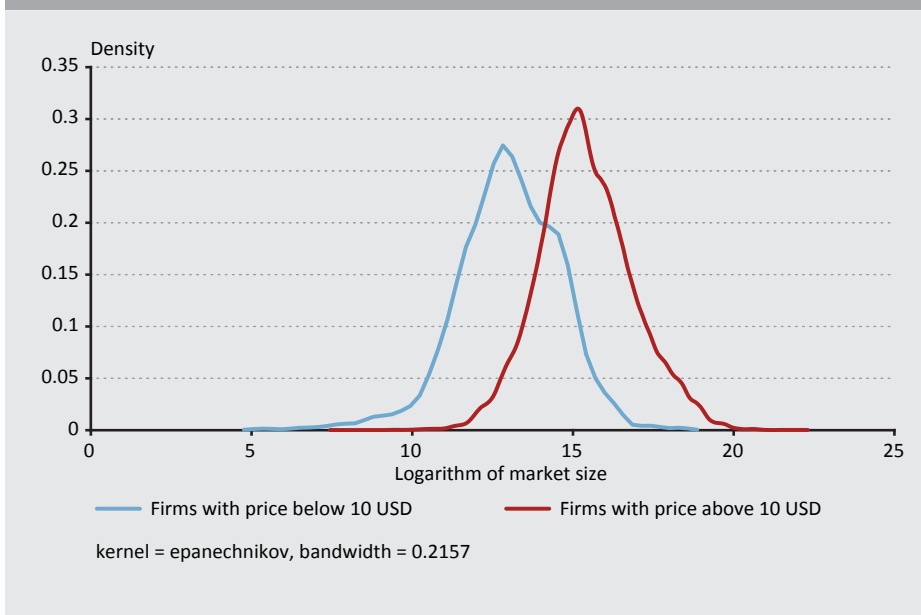
$$BTM_{i,t} = \ln\left(\frac{BVPS_{i,t}}{P_{i,t}}\right) \quad (3)$$

Where $BVPS_{i,t}$ is the book value per share of firm i at the end of the fiscal year to which month t belongs and $P_{i,t}$ is the price of i quoted at the first day of month t .

- *Momentum (Mom)*: the average return on asset i over the last 3 to 12 months.

We followed the procedures proposed by Ince – Porter (2006) with a few modifications and additions for data cleaning. They suggest deleting stocks traded at prices below USD 5. However, our news database represents a news archive of only US equities publicly traded on NYSE/NASDAQ with a price higher than USD 10 per share. For this reason, we used USD 10 as a threshold instead of USD 5. Figure 2 shows that firms with prices below USD 10 generally have smaller market capitalisation. Before dropping stocks with a price under USD 10, the 5th percentile of market value was USD 12.46. After excluding stocks with a price under USD 10, only 1 per cent of the firms have a market value below USD 12.14. Hence, we did not only delete the cheapest stocks: dropping the stocks with low prices partially tackles the problem of small firms. Otherwise, dropping the smallest 5 per cent of firms is a common way of screening small firms.

Figure 2
Kernel density of firms below and above USD 10



Following *Amihud (2002)*, we also delete the most illiquid stocks every month (first decile based on turnover). Observations with returns equal to zero or over 200 per cent or with a total return index smaller than 1 per cent or with missing variables were also omitted. After these steps, we also ignored months which had less than 50 data points.

After cleaning, the database contains 97,178 observations for 1,983 firms, covering the 122 months between January 2010 to February 2020. On average, we have around 800 observations per month. This data excludes the period of Covid-19 to avoid the discussion of how this special period influences our results. According to the literature, Covid-19 had a strong impact on stocks which represents the special nature of that period for both ESG and non-ESG-related financial questions (e.g. *Demers et al. 2021; Kökény et al. 2022*).

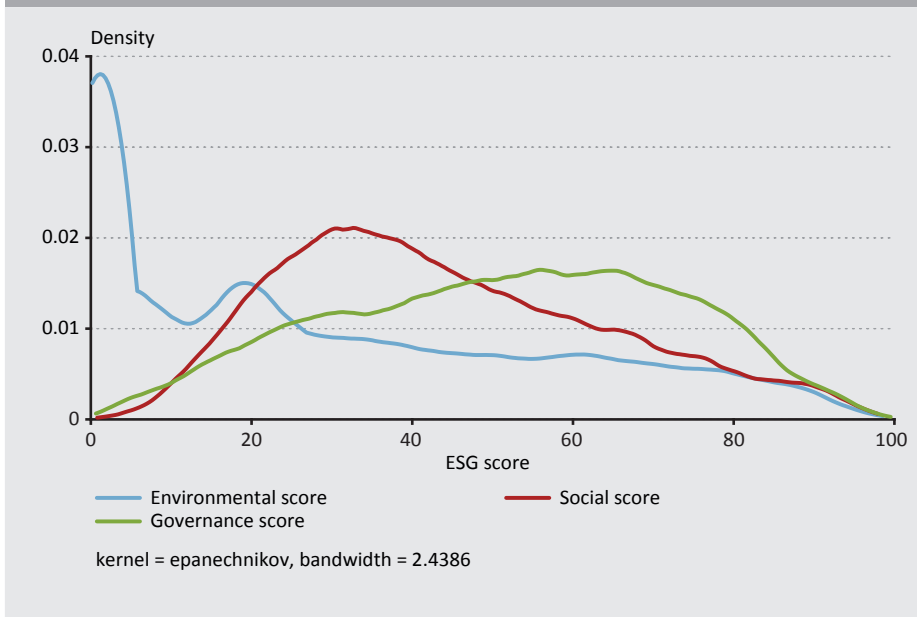
2.3. Descriptive statistics of ESG scores

This section provides descriptive statistics on the ESG scores, with *Table 1* showing the summary statistics. ESG scores take on values between 0 and 100. Generally, the Environmental score is lower relative to Social and Governance scores.

Table 1**Descriptive statistics of ESG scores**

Variable	Mean	Median	St. Dev.	Minimum	Maximum
Environmental Score	28.59	21.14	27.34	0	98.55
Social Score	43.98	40.685	20.23	0.6	97.95
Governance Score	51.27	52.6	21.4	0.44	99.54

Figure 3 presents that stocks in our sample generally have lower environmental scores and higher governance scores. A large share of the companies has an environmental score close to zero, which may bias the results. Around 20 points there is a second spike in the distribution of environmental scores. This might be incorrect data. However, we could not identify any data cleaning step after which the suspicious spike disappears.

Figure 3**Distribution of ESG scores**

3. Methodology

This section presents the applied methodologies in detail. The first step to measure the impact of the news on financial markets is to generate a time series that quantifies the green intensity. For this purpose, we use the topic models. The steps for topic modelling are detailed in *Section 3.1*, while *Section 3.2* presents the methodology of the Fama-MacBeth regression. There are several empirical

approaches for investigating the relationship between a characteristic such as an ESG score and expected returns (Mérő *et al.* 2020). The standard portfolio-based approaches (e.g. *Neszveda – Vágó 2021; Neszveda – Simon 2021*) or alternative portfolio-based approaches (e.g. *Fain – Naffa 2019; Naffa – Fain 2022*) have the advantage that they do not assume any linear relation and reduce noise, but compared to Fama-MacBeth regressions these approaches have the disadvantage of losing information by creating portfolios. Furthermore, Fama-MacBeth regressions are stricter in controlling for other known characteristics related to expected returns. Consequently, we apply Fama-MacBeth regressions to test our main hypothesis.

3.1. Structural topic model

The topic model is a commonly used, unsupervised machine learning tool that identifies topics based on the pattern of occurrence of words in text data. The most common algorithm among topic models is the Latent Dirichlet Allocation (LDA). The general concept of this is that every topic occurs in a certain proportion in every text, and every topic is a mixture of words (Blei *et al.* 2003).

LDA is a statistical method to find the associated words, sort them into topics, and estimate the proportions of these topics in the documents (news articles in our case). As a result, we can derive several coefficients for each word, which shows the probability of that coming from a given topic.

The number of topics is the only tuned hyperparameter of the model. Too few topics will produce overly broad results and it is impossible to interpret them, while choosing too many topics will result in the “over-clustering” of a corpus into many small, highly similar topics (Greene *et al.* 2014, p. 498). Similar to clustering methods, there are rules to determine the optimal value of this parameter, for example, semantic cohesion and exclusivity (Bischof – Airoldi 2012), but we apply a different framework, because of the interest in one topic.

We estimate the model with several numbers of topics (from 2 to 30 for each even value) and use the one with the smallest number of topics and containing one in an identifiable way related to the environment. The proportion of that topic in the news on a given day can be used to extend the general Fama-MacBeth regression. The reason is that if we use a model not having enough topics, then the environmental topic will be mixed with something else, while a model with too many topics would require additional effort to correctly aggregate the environmental topics (we did not observe the appearance of multiple environmental topics with a higher number of topics).

The general approach of Structural Topic Modelling (STM) is very similar to the LDA, but it also uses metadata information on the documents. The proportion in which each topic contributes to a document is called topic prevalence. With the structural

topic model, the prevalence can vary according to the metadata. This causes that, in contrast to the LDA case, the expected topic proportion is not equal for each topic.

Following *Cerchiello – Nicola (2018)* and *Dybowski – Kempa (2020)*, we extend the topic model with *time covariates*, which means that the prevalence of the topic varies over time, and *some topics may go out of fashion while others start trending*. A spline function is used on the number of days since the first article, to ensure that non-linear effects can also be captured. This also results in the prevalence of topics possibly not being equal.

The first step to run STM is to assign a corpus (a collection of words in an ordered form) from the investigated text. Here, we follow the framework of *Roberts et al. (2019)*. As it is a commonly used approach, we omit numbers, stop words (e.g. “and”, “or”, “the”) from the corpus and stem the words (remove the “s”, “tion” and others from the end of the word). We estimate the model with several numbers of topics, from 2 to 30 for each even value, since the runtime is still two days.³

To identify each topic, we use the words that are most likely to come from that topic. The probability of the occurrence of topics in each article is estimated by the model, and the occurrence of each topic about the environment is averaged for a given monthly intensity of the green topic (γ). At this point, we have two options to aggregate: (1) calculating the monthly average disregarding the length of the texts assuming that an article does not carry more weight just because it is longer or weighting the scores by the number of words, or (2) using weighted average. For robustness check, we calculate the Fama-MacBeth model with both frameworks.

3.2. Fama-MacBeth regressions

We first run the following time series regression on the future return for all stocks to obtain each stock's $i \in \{1, \dots, I\}$ exposure to the $m \in \{Beta, MV, BTM, Mom, ES, SS, GS\}$ variables.

$$r_{i,t+1} = \alpha_t + \beta_{t,F_1} F_{1,t} + \dots + \beta_{t,F_m} F_{m,t} + \varepsilon_{i,t} \quad (4)$$

Finally, we test whether the average of the estimated betas is equal to 0 for a given factor in a given time period. A significant average beta suggests that the corresponding factor can predict future returns in a given time period.

4. Results

In the discussion of the results, we keep the same order as for the methodological description, but the descriptive statistics of the green intensity index generated by the topic model are presented between the two model frameworks.

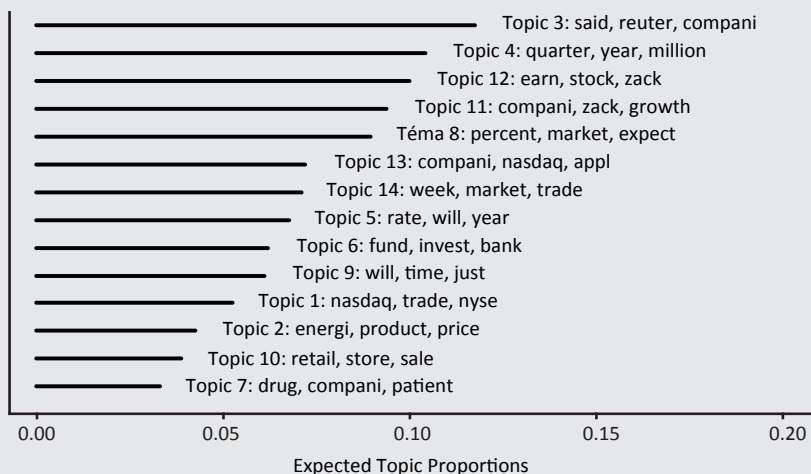
³ CPU: Apple M1 Pro (10 cores), RAM: 17.2 GB.

4.1. Structural Topic Model

Running the structural topic model with a different number of topics, we found that an environmental topic appears first if we categorise the texts into 14 topics.⁴ The topic proportions are demonstrated in *Figure 4*, the environment-related topic is the second one. In support, *Figure 5* shows the words that are most likely to have been generated from Topic 2.

Figure 4

Estimated topic proportions



Note: The word stems are displayed in the figure, because the applied modelling procedures are based on them.

Examining the texts containing the highest estimated proportion of Topic 2 by the model is also apt to confirm our assertion that the occurrence of these topics is a good measure of the intensity of the green topic in the news. Three of these are listed in the following as examples.

“Exxon Mobil NYSE XOM says it is restarting its 560K bbl day Baytown Tex refinery second largest in the US six days after it was shut because of heavy rain from Hurricane Harvey Phillips 66 NYSE PSX says it is preparing to resume operations at its Sweeny refinery and its Beaumont terminal in Texas its Pasadena refined products terminal is resuming truck loading for gasoline this afternoon while operations at its Gulf Coast fractionation plant in Mont Belvieu are suspended Also Occidental Petroleum NYSE OXY has loaded and shipped its first crude oil cargo from its Western Gulf Coast terminal at the Port of Corpus Christi since Harvey”

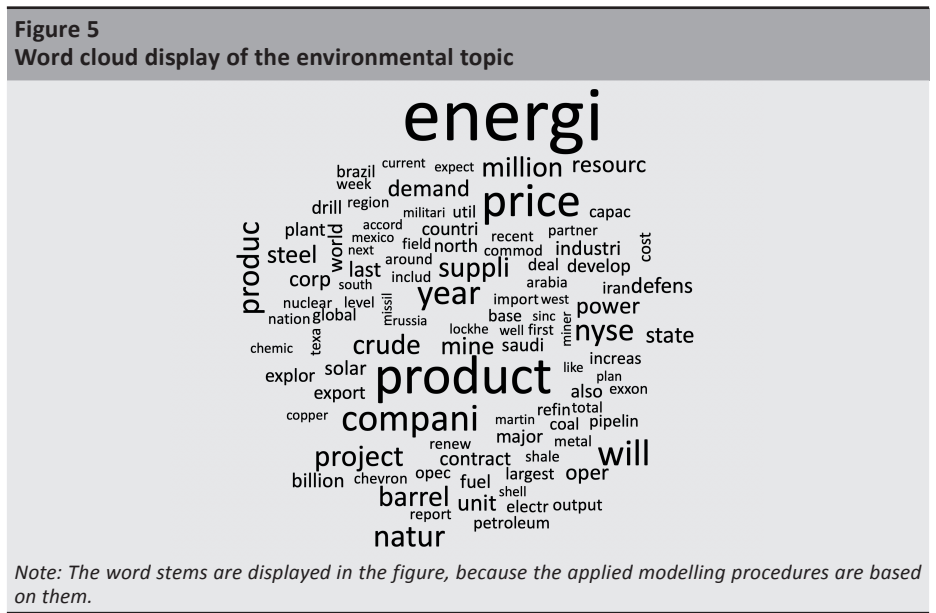
“Exxon Mobil NYSE XOM has made its seventh major oil discovery in the Stabroek block offshore Guyana following drilling at the Pacora 1 exploration well partner Hess NYSE

⁴ The outcome of the other models can be found in the related GitHub repository: <https://github.com/MarcellGranat/green-finance-news/blob/main/result.md>.

HES reveals Pacora resources will be integrated into the third phase of development at the Guyana project helping bring production to more than 500K bbl day of oil Hess says The Pacora 1 well discovery adds to previous world class discoveries at Liza Payara Snoek Liza Deep Turbot and Ranger 1 which are estimated to total more than 3 2B recoverable oil equivalent barrels XOM is operator of the 6 6M acre Stabroek block and holds a 45 while Hess owns 30 and Cnooc s NYSE CEO Nexen has 25”

“Petrobras NYSE PBR subsidiary In Bolivia and the country s state run YPFB Chaco have signed a 1 2B agreement to explore two natural gas fields with potential reserves of 4T of the Bolivian government says The fields are Astillero and San Telmo in southern Bolivia YPFB has a 40 stake in San Telmo and PBR has 60 while PBR owns 40 in Astillero and YPFB has 60 Both fields are expected to begin producing gas in 2022”

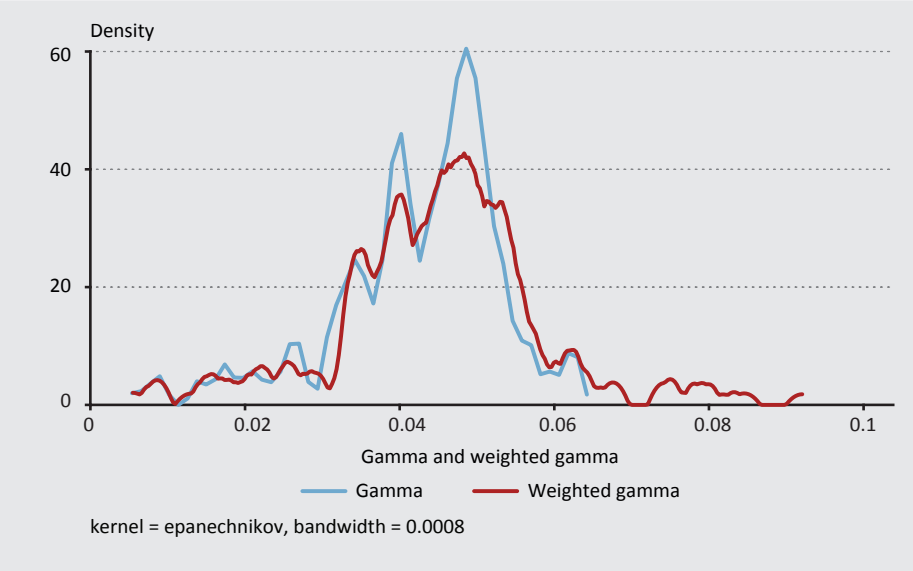
It can be seen from the examples that the topics discussed contain terms mainly *related to the energy sector*: energy, nature, barrel, nuclear, chemical, and mine. We use the proportion of this topic in the articles to proxy the media’s attention to environmental issues.



4.2. Descriptive statistics of the prevalence of environmental topics

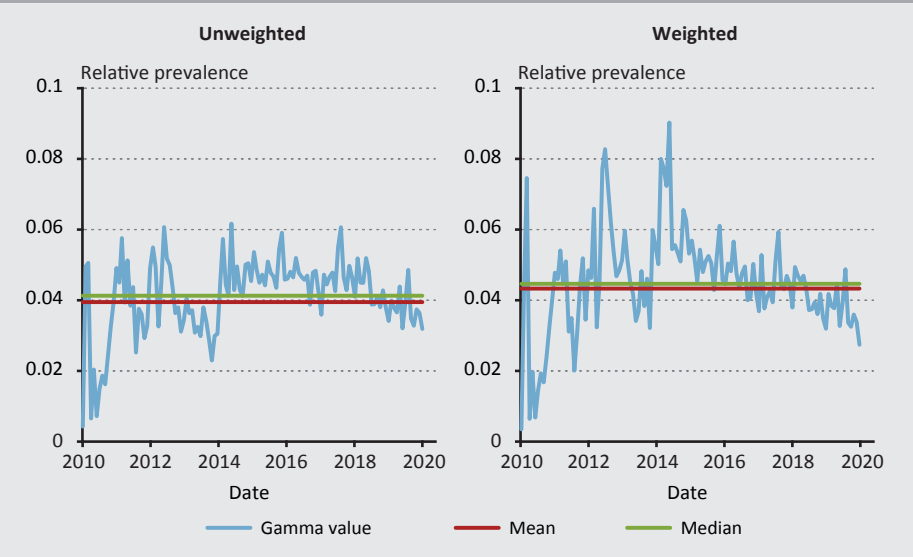
The following subsection describes the time series describing the intensity of the environmental topic generated by STM. As mentioned, we aggregate the prevalence in articles to the monthly level weighting by the number of words in the texts and unweighting. *Figure 6* shows that unweighted and weighted monthly prevalence values are concentrated around 0.032, 0.039 and 0.045, but the distribution of weighted values is smoother.

Figure 6
Distribution of unweighted and weighted prevalence of green topic



In *Figure 7*, we see the time series plot of unweighted and weighted prevalence between 2010 and 2020. The mean and median values are the thresholds used to define high and low-intensity periods, leading to four possible formalisations of intensity.

Figure 7
Monthly aggregated prevalence (January 2010 – February 2022)



4.3. Fama-MacBeth regressions

In this section, we examine whether the effect of ESG scores is different when environmental issues receive more media coverage. To check this, we run Fama-MacBeth regressions (*Fama – MacBeth 1973*) on different subsamples. First, we run regressions using observations in months when the environment received high media coverage. These months are referred to as intensive periods. Then, we use observation in months when environmental issues were less intensively presented in the media. These months are referred to as non-intensive periods. We hypothesise that ESG scores will have a higher effect on future returns in periods of high intensity media coverage. We control for the variables in the Carhart four-factor model (*Carhart 1997*), namely the respective market beta, size, book-to-market ratio and momentum.

We define intensive periods of environmental media coverage in the following four ways:

1. $\gamma > \gamma_{mean}$: *Figure 4* shows which 73 months are identified as intensive periods based on this criterion (*Table 2: 2nd column*). The remaining 49 months are the non-intensive periods (*Table 2: 3rd column*).
2. *weighted* $\gamma > \textit{weighted } \gamma_{mean}$: *Figure 4* shows which 67 months are identified as intensive periods based on this criterion (*4th column*). The remaining 55 months are the non-intensive periods (*5th column*).
3. $\gamma > \gamma_{median}$: *Figure 5* shows which 72 months are identified as intensive periods based on this criterion (*6th column*). The remaining 50 months are the non-intensive periods (*7th column*).
4. *weighted* $\gamma > \textit{weighted } \gamma_{median}$: *Figure 5* shows which 67 months are identified as intensive periods based on this criterion (*8th column*). The remaining 55 months are the non-intensive periods (*9th column*).

We summarise our results in *Table 2. Column 1* shows that in the period January 2010 to February 2020 *Environmental Score* had no significant impact on stock returns. However, *Social Score* was associated with an average monthly return of 0.0062 per cent (Newey-West t-statistic: 2.27), while *Governance Score* decreased future returns c.p. on average by –0.004 per cent (Newey-West t-statistic: –2.45).

These results suggest that stocks with a higher *Social Score* outperform firms with lower social indicator values, while firms with higher *Governance Score* have lower future returns than those with a lower score. The direction of the relationship between these scores and future returns does not change between intensive and non-intensive periods.

Environment and *Governance Scores* have a negative impact on future returns, while *Social Score* has a positive impact. This suggests that different ESG scores are considered differently by investors. Also, these results might be driven by investors' different perceptions of different industries. To analyse this question, one should group stocks into sectors and focus on how ESG scores change among industries. However, this is beyond the scope of this research. Overall, we obtain a rather mixed picture, which is in line with the results of *Cao et al. (2020)*: they find ESG scores to have significant impacts on specific industries, and no effect on others.

Even in some cases where ESG was significant, in general, it causes a decrease in the expected return. This suggests that investors do not value sustainability to the extent that it constrains companies or distracts them from their focus on customers and shareholders. Our results are similar to previous studies focusing on the US market that used the Fama-MacBeth regression methodology (*Timár 2021*). Since ESG scores take on values between 0 and 100, which is significantly different from what characterises the other factors of the regression, comparison of the relative sizes of the coefficients is not feasible.

We find that in periods of intensive environmental media coverage *Social* and *Governance Scores* significantly affect future returns, while in low-intensity periods they have no explanatory power. These findings are robust to different definitions of intensive and non-intensive periods. Moreover, we find that – using the median as a threshold to define high-intensity periods – all of the ESG pillar scores significantly predict future returns.

Our findings also show that in intensive periods the average of the coefficients of *Environmental Score* is -0.0049 , while in non-intensive periods the average of the beta is -0.0017 . The effect is almost three times larger in intensive periods relative to non-intensive months.

These results imply that investors pay attention to ESG scores when environmental issues receive high media coverage. However, in months when investors are less confronted with environmental problems in the press, they do not take these issues into account in their investments.

Table 2
Fama-MacBeth regressions

Months in regression	Benchmark (1) Every month	Identifying intensive periods based on y			Identifying intensive periods based on weighted y				
		Intensive (2) $y > y_{\text{mean}}$	Non-Intensive (3) $y \leq y_{\text{mean}}$	Intensive (6) $y > y_{\text{median}}$	Non-Intensive (7) $y \leq y_{\text{median}}$	Intensive (4) $wy > w.y_{\text{mean}}$	Non-Intensive (5) $wy \leq w.y_{\text{mean}}$	Intensive (8) $wy > w.y_{\text{median}}$	Non-Intensive (9) $y \leq w.y_{\text{median}}$
Beta (A)									
Average return	0.000	0.002	-0.003	0.003	-0.005	0.003	-0.005	0.004	-0.005
NW t-statistics	-0.110	0.350	-0.550	0.600	-0.860	0.630	-0.870	0.740	-0.980
NW p-value	0.911	0.728	0.582	0.549	0.396	0.530	0.389	0.459	0.331
MV (B)									
Average return	0.000	0.000	0.001	0.000	0.001	0.000	0.001	0.000	0.001
NW t-statistics	0.350	-0.690	1.010	-0.610	0.840	-0.520	0.780	-0.200	0.500
NW p-value	0.723	0.492	0.316	0.544	0.407	0.604	0.439	0.846	0.616
BTM (C)									
Average return	-0.001	-0.001	-0.001	-0.001	-0.001	0.000	-0.002	0.000	-0.001
NW t-statistics	-1.400	-0.630	-1.520	-0.590	-1.370	-0.410	-1.570	-0.560	-1.470
NW p-value	0.165	0.532	0.136	0.554	0.177	0.682	0.123	0.577	0.148
Mom (D)									
Average return	0.001	0.004	-0.004	0.005	-0.004	0.003	-0.003	0.005	-0.004
NW t-statistics	0.400	0.930	-0.990	1.100	-1.230	0.830	-0.670	1.230	-1.040
NW p-value	0.690	0.358	0.325	0.275	0.223	0.408	0.504	0.224	0.304
ES (E)									
Average return	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
NW t-statistics	-1.400	-1.590	-0.690	-1.750	-0.540	-1.430	-0.420	-1.810	-0.110
NW p-value	0.164	0.116	0.496	0.086*	0.589	0.157	0.679	0.075*	0.911
SS (F)									
Average return	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
NW t-statistics	2.270	2.500	0.710	3.040	0.540	2.110	1.050	2.370	0.810
NW p-value	0.025**	0.015**	0.482	0.003***	0.592	0.038**	0.301	0.021**	0.422
GS (G)									
Average return	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
NW t-statistics	-2.450	-1.940	-1.590	-1.880	-1.600	-2.110	-1.570	-2.230	-1.480
NW p-value	0.016***	0.057*	0.118	0.065*	0.115	0.039***	0.122	0.029***	0.145

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. In Column 1, we run FM regressions using every observation in months between January 2010 to February 2020. In Columns 2 to 9, we run FM regressions using observations from months that meet given criteria. For instance, in Column 2, we use months in which y was higher than y mean. This way there are gaps in the time series in Columns 2 to 9. Average return shows the estimated coefficients, NW-t-statistics shows Newey-West t-statistics with 12 months lag, while NW p-value shows corresponding p-value. Beta is monthly beta; MV is the logarithm of the stocks' market capitalisation; BTM is the logarithm of the firms' book-to-market ratio, while Mom is the momentum; the average return over last year; ES, SS and GS show weighted average relative ESG ratings of a company.

5. Conclusion

This paper examined the relationship between the effect of ESG scores on future returns and the intensity of media coverage of environmental issues. We assessed the importance of environmental issues in the media with a machine learning-based Structural Topic Modelling (STM) methodology. We subsequently ran Fama-MacBeth regressions (Fama – MacBeth 1973) using observations in months when environmental issues received higher-than-average (lower-than-average) media coverage. We found that in intensive topic periods (using the median as a threshold) ESG scores negatively and significantly affect future returns, while in low-intensity periods they have no explanatory power. These results suggest that investors pay less attention to firms' attitudes towards sustainability when they see fewer articles related to environmental issues. However, when environmental issues receive high media coverage investors do consider ESG scores. Generally, we find a negative relationship between *Environmental score* and future return, irrespective of the media's attention to climate change. This contradicts the idea that investors value sustainability. However, it is important to mention that we did not consider that the impact of the ESG scores might differ between industries and ignored the fact that retail traders may assess ESG scores differently than institutional investors. Hence, our results are not intended to measure the magnitude of the effect, but how the effect varies with media attention.

Our research has three important limitations. We adopt gamma's and weighted gamma's statistical measures of central tendency (mean, median) to define intensive and non-intensive periods. We used monthly gamma and weighted gamma values over the period from October 2008 to February 2020 to assess the mean and median. However, this information became available only as time progressed. Thus, we calculate media intensity in period t using information that was not yet available in that month. This method was necessary because we did not have any prior knowledge about which gamma values should be considered relatively high or low.

Second, we do not answer which factors are causing the phenomenon we have found. Previous research found that capital market anomalies are amplified by the presence of small investors. Previous studies (e.g. Csillag – Neszveda 2020; Choi et al. 2020) also find that retail investors' (not institutional investors') ESG trading patterns are exposed to extreme weather in their location. This study did not examine whether our finding was led by increased retail investor attention to global warming in periods of intensive media coverage.

Also, this study did not seek to uncover what causes the deviation in media coverage of environmental issues. However, to truly understand the connection between our *gamma* and ESG scores one must answer that question. Based on our results this might be a reasonable next step. The paper of Choi et al. (2020) suggests that

extreme weather might grab press attention, resulting in more intensive media coverage of environmental issues. However, as our news data source is investment specific it might be less exposed to this effect. Also, other events may affect the intensity of the environmental topic, such as energy, environmental regulation and weather-related events. Moreover, further research could analyse how media intensity affects ESG scores in different sectors. Finally, another interesting question could be whether media intensity differently impacts low-emission “green” stocks and high-emission “brown” stocks.

References

- Akerlof, K. – Maibach, E.W. – Fitzgerald, D. – Cedenó, A.Y. – Neuman, A. (2013): *Do people “personally experience” global warming, and if so how, and does it matter?* Global Environmental Change, 23(1): 81–91. <https://doi.org/10.1016/j.gloenvcha.2012.07.006>
- Amihud, Y. (2002): *Illiquidity and stock returns: Cross-section and time-series effects*. Journal of Financial Markets, 5(1): 31–56. [https://doi.org/10.1016/S1386-4181\(01\)00024-6](https://doi.org/10.1016/S1386-4181(01)00024-6)
- Auer, B.R. – Schuhmacher, F. (2016): *Do socially (ir) responsible investments pay? New evidence from international ESG data*. The Quarterly Review of Economics and Finance, 59(February): 51–62. <https://doi.org/10.1016/j.qref.2015.07.002>
- Bischof, J. – Airoldi, E.M. (2012): *Summarizing topical content with word frequency and exclusivity*. Proceedings of the 29th International Conference on Machine Learning (ICML-12), 201–208. <https://icml.cc/Conferences/2012/papers/113.pdf>
- Blei, D.M. – Ng, A.Y. – Jordan, M.I. (2003): *Latent Dirichlet Allocation*. Journal of Machine Learning Research, 3(Jan): 993–1022. <https://www.jmlr.org/papers/volume3/blei03a/blei03a.pdf>
- Broomell, S.B. – Budescu, D.V. – Por, H.-H. (2015): *Personal experience with climate change predicts intentions to act*. Global Environmental Change, 32(May): 67–73. <https://doi.org/10.1016/j.gloenvcha.2015.03.001>
- Buallay, A. (2019): *Is sustainability reporting (ESG) associated with performance? Evidence from the European banking sector*. Management of Environmental Quality, 30(1): 98–115. <https://doi.org/10.1108/MEQ-12-2017-0149>
- Cao, J. – Titman, S. – Zhan, X. – Zhang, W. (2020): *ESG preference, institutional trading, and stock return patterns*. Working Paper 28156, National Bureau of Economic Research. <https://doi.org/10.3386/w28156>
- Carhart, M. M. (1997): *On Persistence in Mutual Fund Performance*. The Journal of Finance, 52(1): 57–82. <https://doi.org/10.1111/j.1540-6261.1997.tb03808.x>

- Cerchiello, P – Nicola, G. (2018): *Assessing News Contagion in Finance*. *Econometrics*, 6(1): 5. <https://doi.org/10.3390/econometrics6010005>
- Choi, D. – Gao, Z. – Jiang, W. (2020): *Attention to Global Warming*. *The Review of Financial Studies*, 33(3): 1112–1145. <https://doi.org/10.1093/rfs/hhz086>
- Csillag, B. – Neszveda, G. (2020): *A gazdasági várakozások hatása a tőzsdei momentumstratégiára*. *Közgazdasági Szemle*, 67(11): 1093–1111. <https://doi.org/10.18414/KSZ.2020.11.1093>
- DellaVigna, S. – Pollet, J.M. (2009): *Investor inattention and Friday earnings announcements*. *The Journal of Finance*, 64(2): 709–749. <https://doi.org/10.1111/j.1540-6261.2009.01447.x>
- Demers, E. – Hendrikse, J. – Joos, P. – Lev, B. (2021): *ESG did not immunize stocks during the COVID-19 crisis, but investments in intangible assets did*. *Journal of Business Finance & Accounting*, 48(3–4): 433–462. <https://doi.org/10.1111/jbfa.12523>
- Dybowski, T.P. – Kempa, B. (2020): *The European Central Bank's monetary pillar after the financial crisis*. *Journal of Banking & Finance*, 121(December): 105965. <https://doi.org/10.1016/j.jbankfin.2020.105965>
- Fain, M. – Naffa, H. (2019): *Performance Measurement of Active Investment Strategies Using Pure Factor Portfolios*. *Financial and Economic Review*, 18(2). 52–86. <https://doi.org/10.33893/FER.18.2.5286>
- Fama, E.F. – MacBeth, J.D. (1973): *Risk, Return, and Equilibrium: Empirical Tests*. *Journal of Political Economy*, 81(3): 607–636. <https://doi.org/10.1086/260061>
- Flammer, C. (2013): *Corporate Social Responsibility and Shareholder Reaction: The Environmental Awareness of Investors*. *Academy of Management Journal*, 56(3): 758–781. <https://doi.org/10.5465/amj.2011.0744>
- Gillan, S. – Hartzell, J.C. – Koch, A. – Starks, L.T. (2010): *Firms' environmental, social and governance (ESG) choices, performance and managerial motivation*. Unpublished working paper. <https://sites.pitt.edu/~awkoch/ESG%20Nov%202010.pdf>
- Greene, D. – O'Callaghan, D. – Cunningham, P. (2014): *How many topics? Stability analysis for topic models*. *Joint European conference on machine learning and knowledge discovery in databases*, pp. 498–513. https://doi.org/10.1007/978-3-662-44848-9_32
- Hirshleifer, D. – Teoh, S.H. (2003): *Limited attention, information disclosure, and financial reporting*. *Journal of Accounting and Economics*, 36(1–3): 337–386. <https://doi.org/10.1016/j.jacceco.2003.10.002>

- Ince, O.S. – Porter, R.B. (2006): *Individual equity return data from Thomson Datastream: Handle with care!* Journal of Financial Research, 29(4): 463–479. <https://doi.org/10.1111/j.1475-6803.2006.00189.x>
- Jain, M. – Sharma, G.D. – Srivastava, M. (2019): *Can Sustainable Investment Yield Better Financial Returns: A Comparative Study of ESG Indices and MSCI Indices.* Risks, 7(1): 15. <https://doi.org/10.3390/risks7010015>
- Kotsantonis, S. – Serafeim, G. (2019): *Four Things No One Will Tell You About ESG Data.* Journal of Applied Corporate Finance, 31(2): 50–58. <https://doi.org/10.1111/jacf.12346>
- Kökény, L. – Kenesei, Z. – Neszveda, G. (2022): *Impact of COVID-19 on different business models of European airlines.* Current Issues in Tourism, 25(3): 458–474. <https://doi.org/10.1080/13683500.2021.1960284>
- Li, Y. – Johnson, E.J. – Zaval, L. (2011): *Local Warming: Daily Temperature Change Influences Belief in Global Warming.* Psychological science, 22(4): 454–459. <https://doi.org/10.1177/0956797611400913>
- Mérő, B. – Nagy, O. – Neszveda, G. (2020): *Testing New Factors in Empirical Asset Pricing.* Szigma, 50(4): 263–281.
- Naffa, H. – Fain, M. (2020): *Performance measurement of ESG-themed megatrend investments in global equity markets using pure factor portfolios methodology.* PloS one, 15(12): e0244225. <https://doi.org/10.1371/journal.pone.0244225>
- Naffa, H. – Fain, M. (2022): *A factor approach to the performance of ESG leaders and laggards.* Finance Research Letters, 44(January): 102073. <https://doi.org/10.1016/j.frl.2021.102073>
- Neszveda, G. (2018): *The Contribution of Thaler to Behavioural Economics.* Financial and Economic Review, 17(1): 153–167. <https://doi.org/10.25201/FER.17.1.153167>
- Neszveda, G. – Simon, P. (2021): *Szezonális, január-hatás és a momentum-stratégia.* Szigma, 52(4): 335–352.
- Neszveda, G. – Vágó, Á. (2021): *A likviditásnyújtás kereskedési stratégiájának hozamvizsgálata a magyar részvénypiacon.* Közgazdasági Szemle, 68(7–8): 794–814. <https://doi.org/10.18414/KSZ.2021.7-8.794>
- Roberts, M.E. – Stewart, B.M. – Tingley, D. (2019): *Stm: An R Package for Structural Topic Models.* Journal of Statistical Software, 91(2): 1–40. <https://doi.org/10.18637/jss.v091.i02>
- Sahut, J.-M. – Pasquini-Descomps, H. (2015): *ESG Impact on Market Performance of Firms: International Evidence.* Management International/International Management/Gestión Internacional, 19(2): 40–63. <https://doi.org/10.7202/1030386ar>

- Schmidt, G. (2015): *Thoughts on 2014 and ongoing temperature trends*. Blog post, Real Climate, 22. <https://edukasi.blog/an-http-www.realclimate.org/index.php/archives/2015/01/thoughts-on-2014-and-ongoing-temperature-trends/>
- Shumway, T. (1997): *The Delisting Bias in CRSP Data*. The Journal of Finance, 52(1): 327–340. <https://doi.org/10.1111/j.1540-6261.1997.tb03818.x>
- Talan, G. – Sharma, G.D. (2019): *Doing Well by Doing Good: A Systematic Review and Research Agenda for Sustainable Investment*. Sustainability, 11(2), 353. <https://doi.org/10.3390/su11020353>
- Timár, B. (2021): *How Does the Market Price Responsible and Sustainable Investments?* Financial and Economic Review, 20(2): 117–147. <http://doi.org/10.33893/FER.20.2.117147>
- Townsend, B. (2020): *From SRI to ESG: The Origins of Socially Responsible and Sustainable Investing*. The Journal of Impact and ESG Investing, 1(1): 10–25. <https://doi.org/10.3905/jesg.2020.1.1.010>