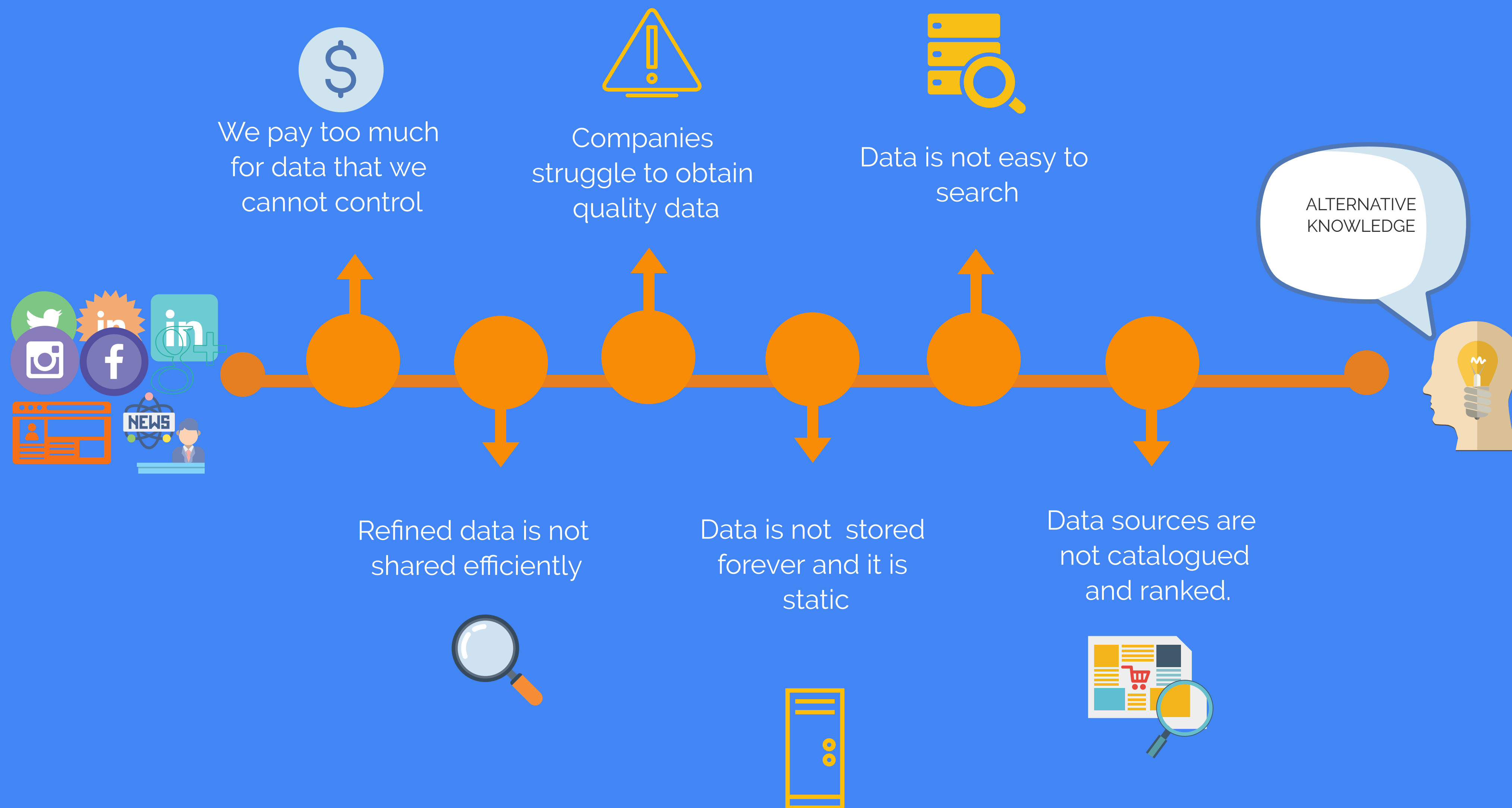
Knowsis ESG

ESG Real-time Sentiment Analysis

Summary

About Knowsis

Knows is is an AI data company extracting values from non-traditional online mediums into quantifiable and actionable output for capital markets and business. Our mission is to develop products and services that enhance the value of social and public information. We *transform varied sources of unstructured data into actionable insights* by applying cutting edge natural language processing techniques and interpretable deep learning algorithms.



The Team

Filippo Lanza, Director

Filippo has more than 20 years of experience in the Hedge Fund industry. Filippo is a strong believer that the perfect team combines Artificial and Human Intelligence to generate a super-data set to inform investment decisions. Over the years Filippo has started and invested in several disruptive and successful tech start ups

Armando Marozzi, Head of Quantitative Research

Armando Marozzi is a post-doctoral researcher at LSE where he carries out research in Macro and Monetary Economics with particular focus on macroeconometric techniques as well as on Natural Language Process (NLP) models applied to macro forecasting. He also worked as a macroeconomist at Medley Global Advisors (MGA), Goldman Sachs and the European Central Bank (ECB).

Tommaso Iaquone, Head of Growth

MSc Finance, MSc Economics. Tommaso has previously worked in the banking industry and managed various development projects in young and emerging start-ups. He is the editor and creator of esgstack.com, the first online space dedicated to resolving ambiguities around ESG. Always willing to create flexible business conditions to guarantee the long-term performance and profitability of young companies and start-ups

Alvise Susmel, CTO

Alvise has previously been CTO of a London-based Hedge Fund, working on different big-data platforms which analyse and process crucial investments' data. Eager to learn new technologies, paradigms and architectures, he also developed a deep interest in teaching.

Nick Oh, NLP Engineer

BSc in Politics and Economics at LSE. Nick is the Founder and Director of Research at SocialScienceAI (socialscience.ai), a Non-Profit Organization with a mission to advance the application of machine and deep learning algorithms in the social sciences space.

Knowsis ESG Model

Context & Methodology

The State of Play on ESG

ESG scope may vary contingent on the legal framework of jurisdictions posing a considerable obstacle for the harmonisation of an ESG definition. This has created the concern among Investment Managers that ESG is going to be more of a moving target rather than a clear measurable input for the investment process, raising many issues and exposing the industry to new risks.

Issues

- ESG data are often available on an annual or quarterly basis, creating a gap between releases
- Existing ESG Scores are backword-looking
- There is no real-time feedback and data is updated infrequently
- Corporate data is self-assessed and biased
- Sentiment computed with generic dictionaries

Our Solutions

- Granular ESG Feedback on the behaviour of Companies and Countries
- Tweets are often forward-looking
- Tweets are transformed into second by second sentiment data
- Tweets are ruthlessly "democratic"
- ESG-specific dictionary and classification

There are two main issues with current products: (i) *low frequency indices* and (ii) *generic sentiment*. ESG-STACK improves on both aspects by providing ***intraday real-time validated score*** at *country*, *sector* and *company* level. We calibrate our model to be tailored to every letter of E-S-G.

System Elements Overview

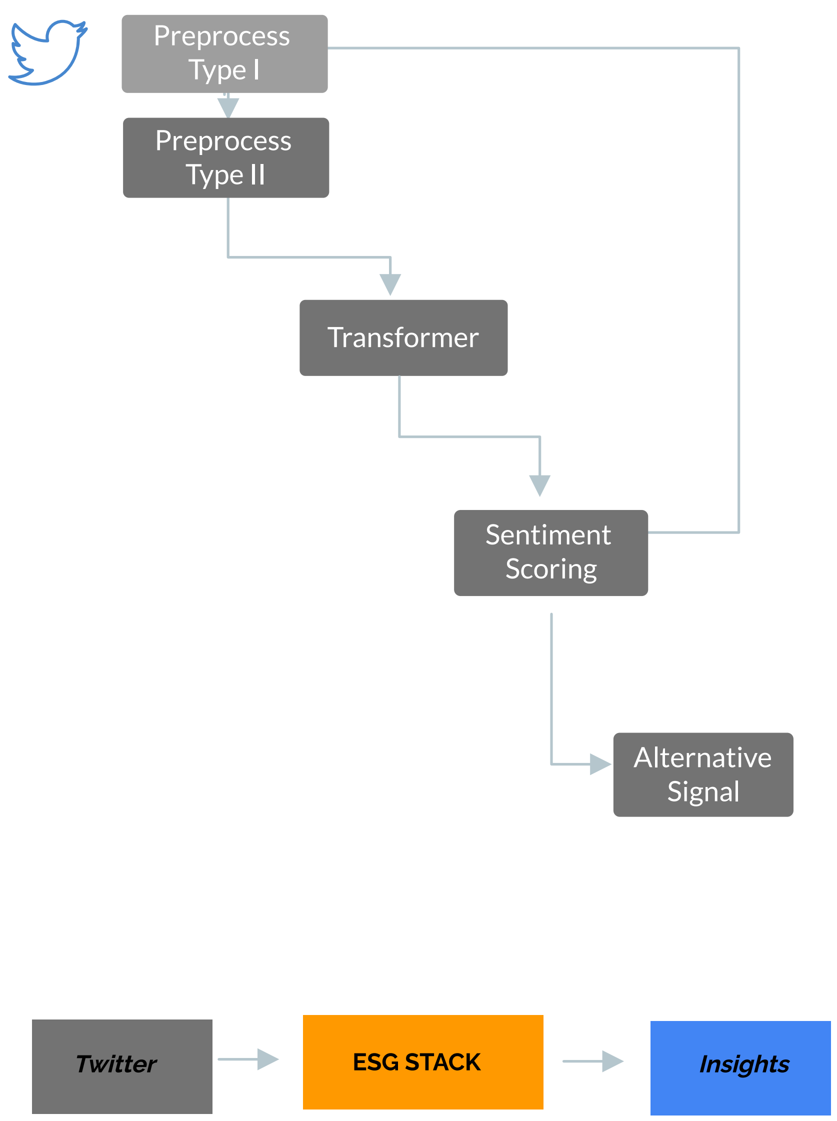
ESG Model

Hybrid Rule-based Transformer Sentiment Analyzer

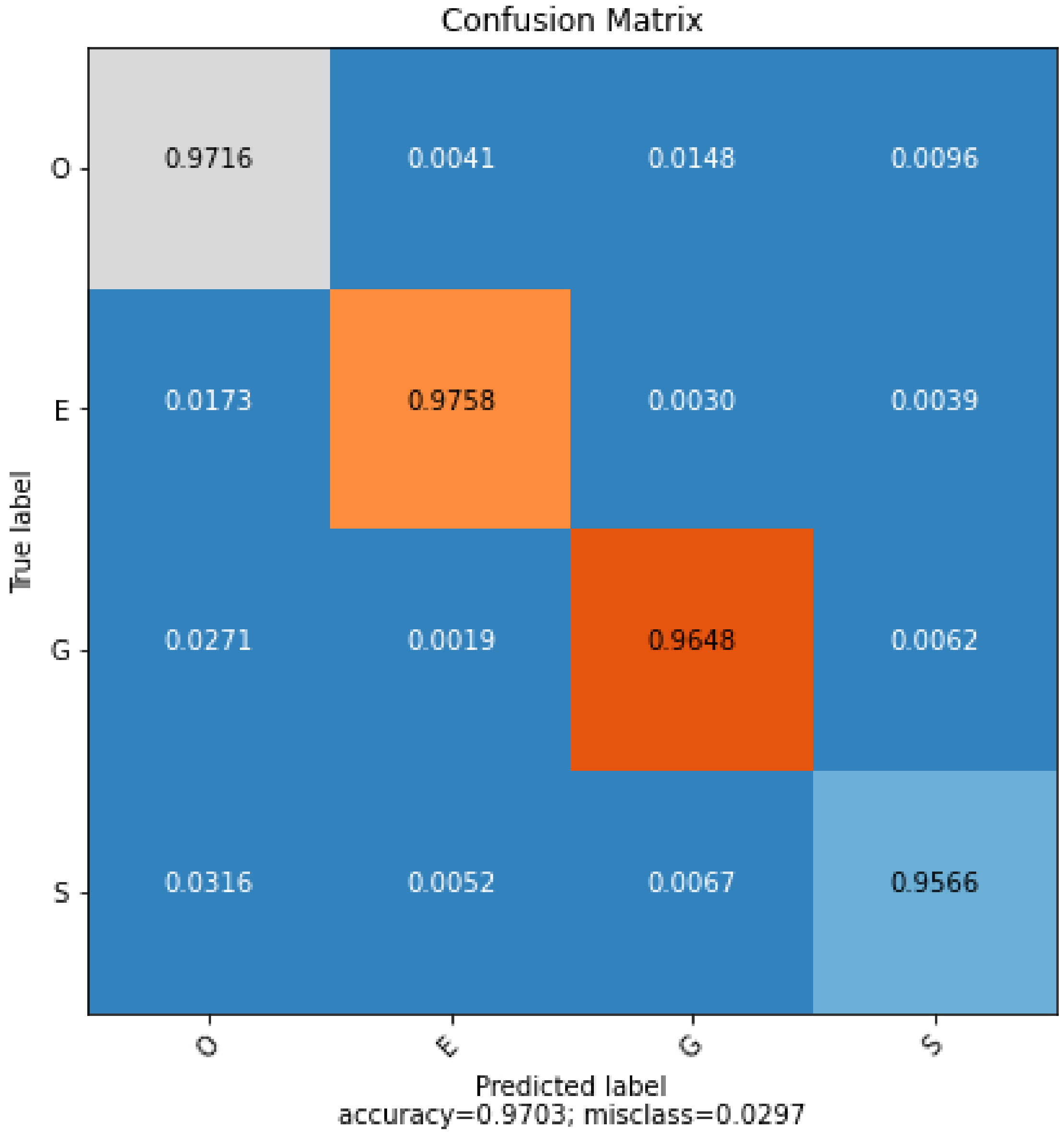
CLIENT

Components of System	Details
Twitter Pre-processor	Two-step pre-processors, each for classification task and for sentiment scoring task: Preparation of tweets with punctuations, emoticons, and capital letters allows a more precise reflection of the "tone" of the message by capturing the valence of the sentiments
Classification Transformer	The transformer model is the backbone of industry-standard NLP models developed by Google AI, Facebook AI, and Open AI (i.e GPT-2, GPT-3, XLNet, BERT, RoBERTa). The custom transformer of the ESG-STACK achieved 97% accuracy in classifying ESG-relevant tweets
Sentiment Scoring	Valence-aware scorings for *15,610* Sentiment Expressions (incl. emoticons) relevant in both context-free (*6,781*) and ESG-specific (*8,829*) tweets
Forecast and Monitor	Valence-aware scorings for *15,610* Sentiment Expressions (incl. emoticons) relevant in both context-free (*6,781*) and ESG-specific (*8,829*) tweets

System Flow



Classification Transformer: Performance



	O	E	G	S
Precision	0.9857	0.9801	0.9353	0.8849*
Accuracy (Recall)	0.9716	0.9757	0.9648	0.9565
F score	0.9786	0.9779	0.9498	0.9193

(*Out of the False Positives of 'S' (i.e tweets that have been wrongly classified as 'S') approximately 80% were 'O'. The cost of falsely predicting 'O' as 'S' are often mitigated as the tweets are asset-related.)

Sentiment Scoring: Interpretable, Explainable and Transparent

ESG-sensitive Sentiment Analysis and Checker

ESG-classified Tweet

Scans **ngrams** in the lexicon dictionary

Identifies **generalizable heuristics** to fine-tune sentiment valence of tweets

(i) *Punctuation* (i.e number of '!' or '?' amplifying sentiment)

(ii) Capitalization (i.e ALL-CAPS emphasizing sentiment)

(iii) *Degree Modifiers* (i.e adverbs intensifying the degree of sentiment)

(iv) *Contrastive conjunctions* (i.e 'but' signalling a shift in sentiment polarity)

(v) *Negations* (i.e 'not' flipping the sentiment polarity)

Sentiment Scoring

[-1,1]

"... if an inherently interpretable model does the job almost as well, companies may opt for that simply because they are easier to explain, defend, and - even for their own developers - to understand and optimize." (David Thogmartin, Director of aiStudio, Deloitte)

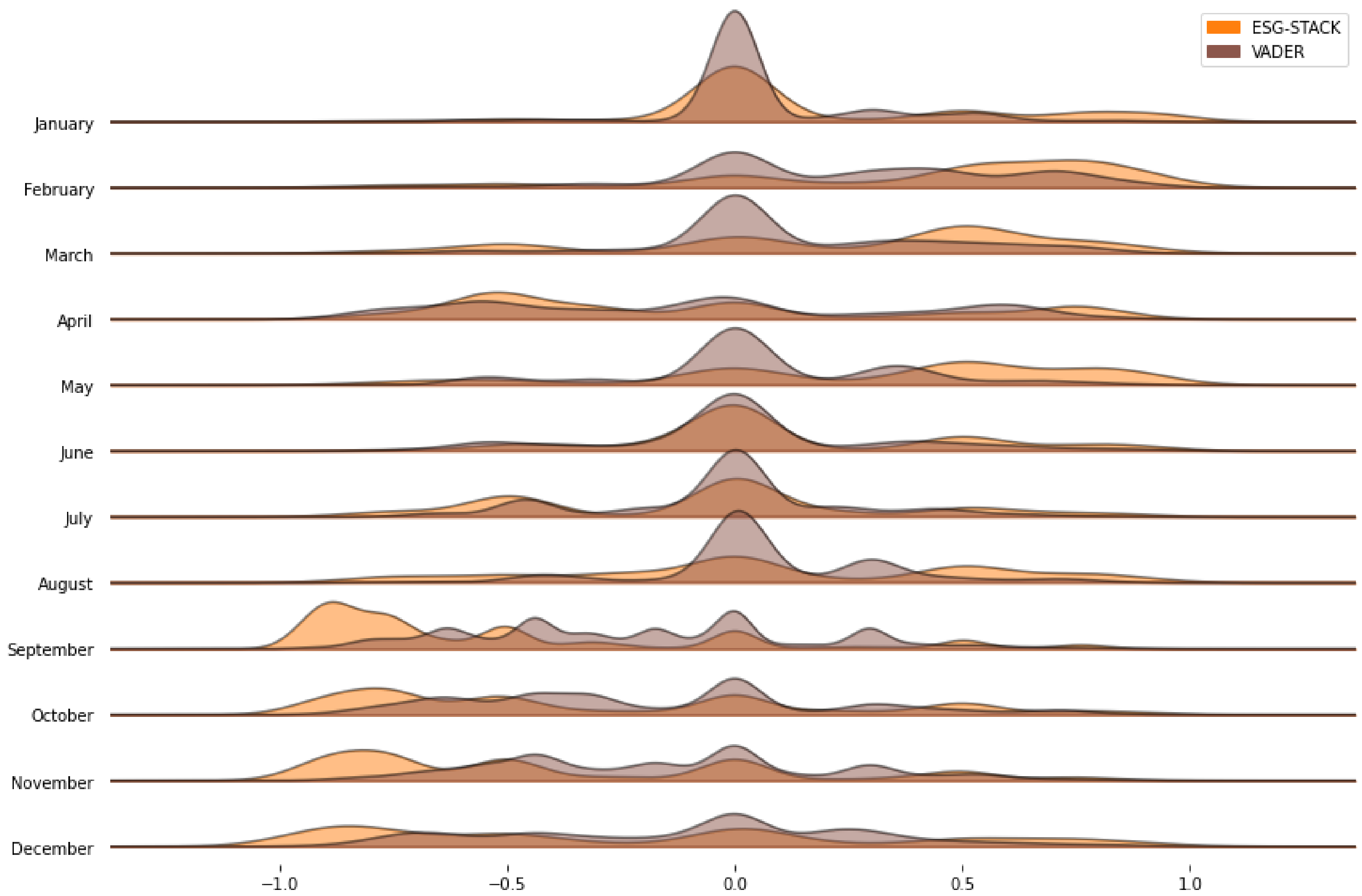
Blackbox
deep-learning

Unreliable & *Unaccountable* Alternative Data

State-of-the-art deep learning (DL) models achieve impressive accuracy when solving NLP tasks. However, DL models are difficult to interpret due to their lack of transparency. Why would you risk making an investment decision based on a black-box system that you cannot fully comprehend?

ESG-STACK instead provides investors with the transparent, reliable and accountable procedure without sacrificing performance for explainability. Our system outperforms the most powerful NLP models such as roBERTa and XLM-roBERTa (see Volkswagen AG case study).

ESG-STACK vs Rule-based Sentiment Analyzer

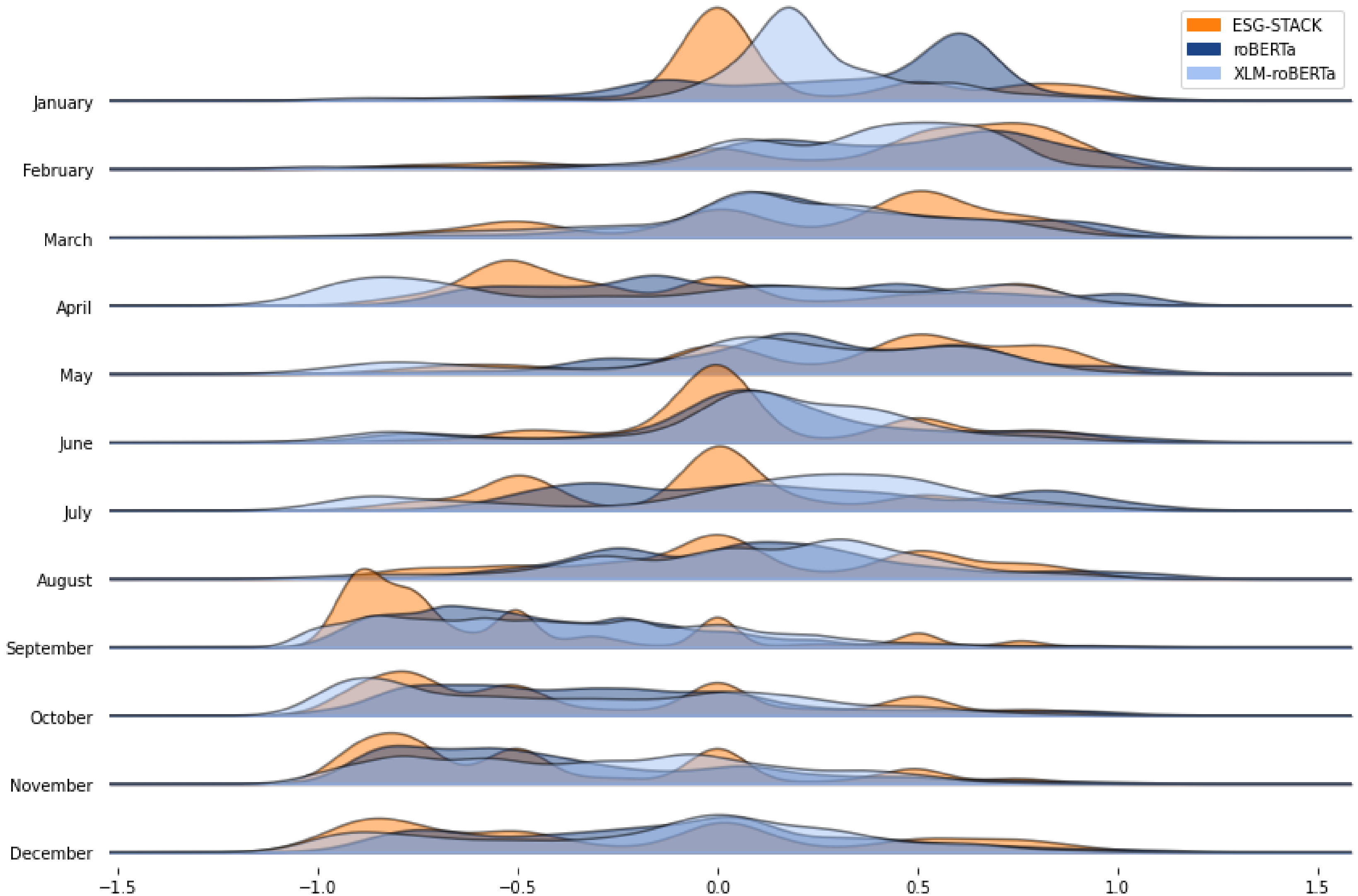


VOW3.DE (Pearson Correlation)		
	corr	**ρ**
ESG-STACK	***0.5979***	***5.195^e-26***
VADER1	0.5090	3.803^e-18
TextBlob2	0.1534	0.01439

[1] see https://github.com/cjhutto/vaderSentiment;
[2] see https://textblob.readthedocs.io/en/dev/

For a discussion of these results please look at Section 2.4 (pp.9-14) of the White Paper: "Modelling ESG with Twitter Data".

ESG-STACK vs State-of-the-Art Transformer Sentiment Analyzer



VOW3.DE *(Pearson Correlation)*		
	corr	**ρ**
ESG-STACK	***0.5979***	***5.195^e-26***
roBERTa1	0.5918	2.115^e-25
XLM- roBERTa2	0.5091	3.709^e-18

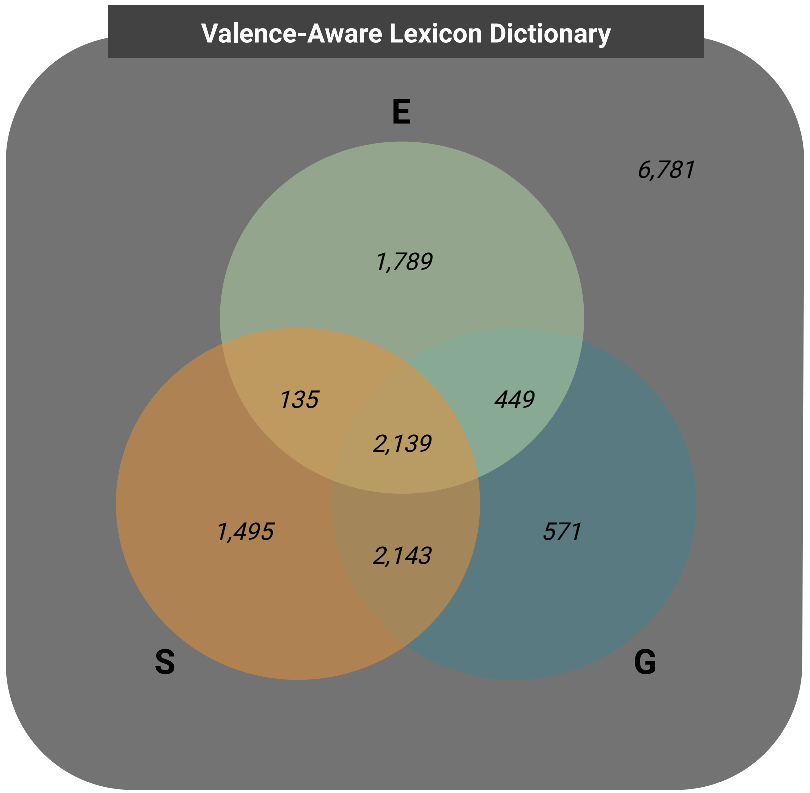
[1] trained on ~58M tweets and finetuned for sentiment analysis (for the details of roBERTa, see https://huggingface.co/cardiffnlp/twitter-roberta-base-sentiment and https://arxiv.org/pdf/2010.12421.pdf);

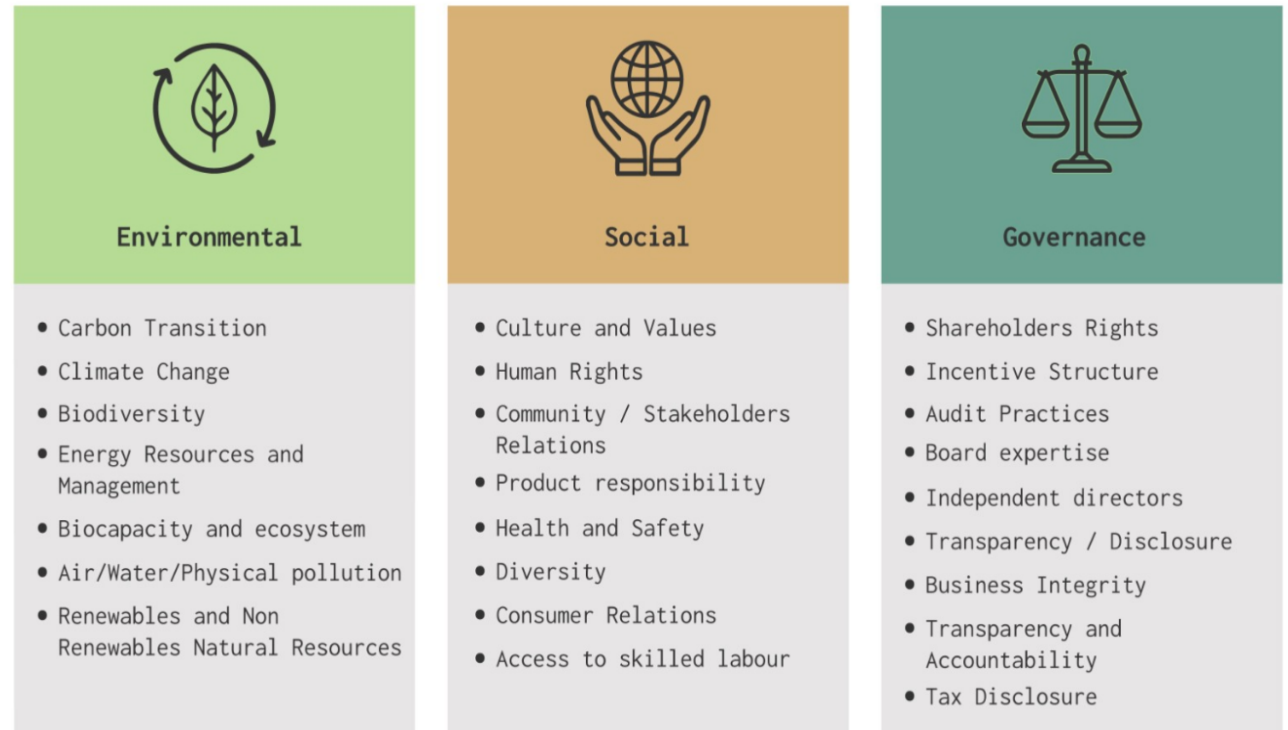
[2] trained on ~198M tweets and finetuned for sentiment analysis (for the details of XLM-roBERTa, see https://arxiv.org/pdf/2104.12250.pdf and https://huggingface.co/cardiffnlp/twitter-xlm-roberta-base-sentiment)

For a discussion of these results please look at Section 2.4(pp.9-14) of the White Paper: "Modelling ESG with Twitter Data".

ESG-specific Lexicon Dictionary: esgstack.com

The lexicon dictionary of ESG-STACK is the ***first and only ESG-sensitive dictionary*** *currently on the market:* It is a dictionary of ***15,610*** **n-grams**[1, 2], cross-validated by *buy-side people* and *ESG experts*, each with valence aware sentiment scorings between [-4,4].





Source: esgstack.com

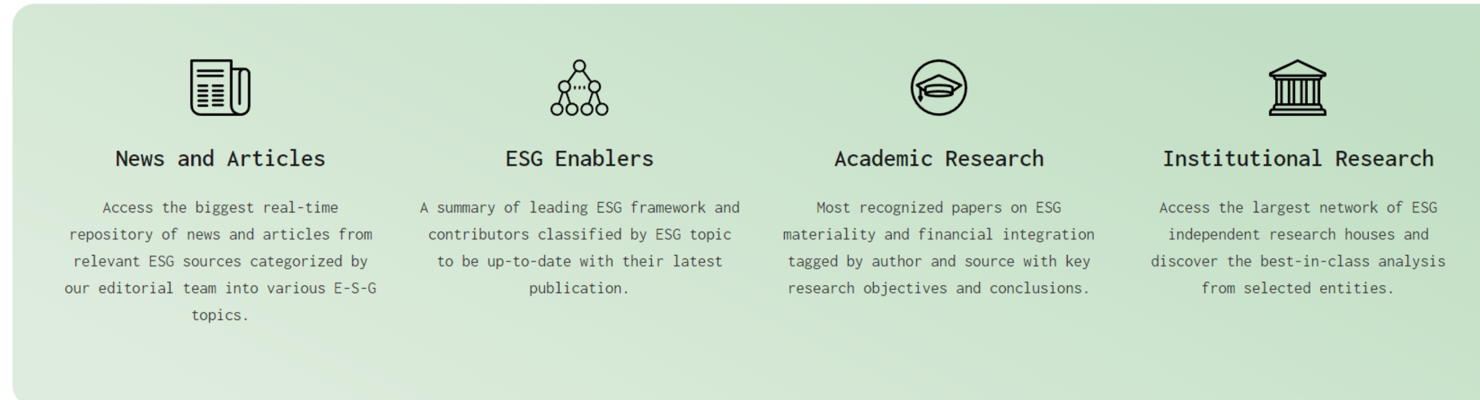
[1] From the gold-standard valence aware sentiment with *7,502* uni-grams of casual sentiment expressions (Hutto and Gilbert 2014), a total of 1,898 expressions have been modified (1,177) and removed (721) to fit in the ESG context

[2] A team of ESG experts extracted a total of *8,829* ESG-specific n-grams from the Twitter – composed of (i) 5,376 unigrams; (ii) 2,520 bigrams, (iii) 790 tri-grams; and (iv) 143 n-grams (4+) – based on Knowsis ESG Mapping

Source: esgstack.com

ESGStack offers a comprehensive, continuously updated smart dictionary on ESG

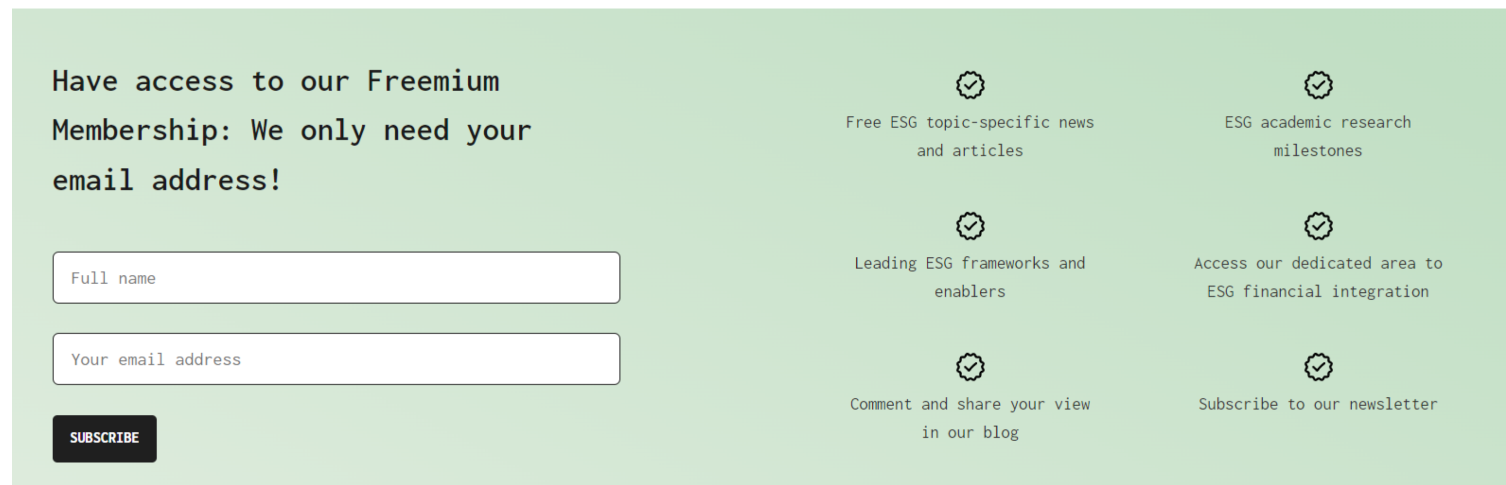
Esgstack is a new project piloted by the **Knowsis** team aiming to simplify access to the best ESG ideas in the street. Our goal is to create a repository of relevant documents that will help our audience to understand more in-depth the complexities around ESG and navigate the multitude of frameworks and initiatives aimed at tackling various ESG topics and taxonomies.



Freemium: Free real-time updated repository of E-S-G letter specific tagged documents. More than 2000+ Academic Papers, News, Articles and ESG Frameworks tagged by ESG topics and subtopics.

Premium: Editorial picks and summarization of most relevant ESG articles and academic papers

Pro: ESG Research House Marketplace. With the help of our, AI find all the relevant institutional research with one click



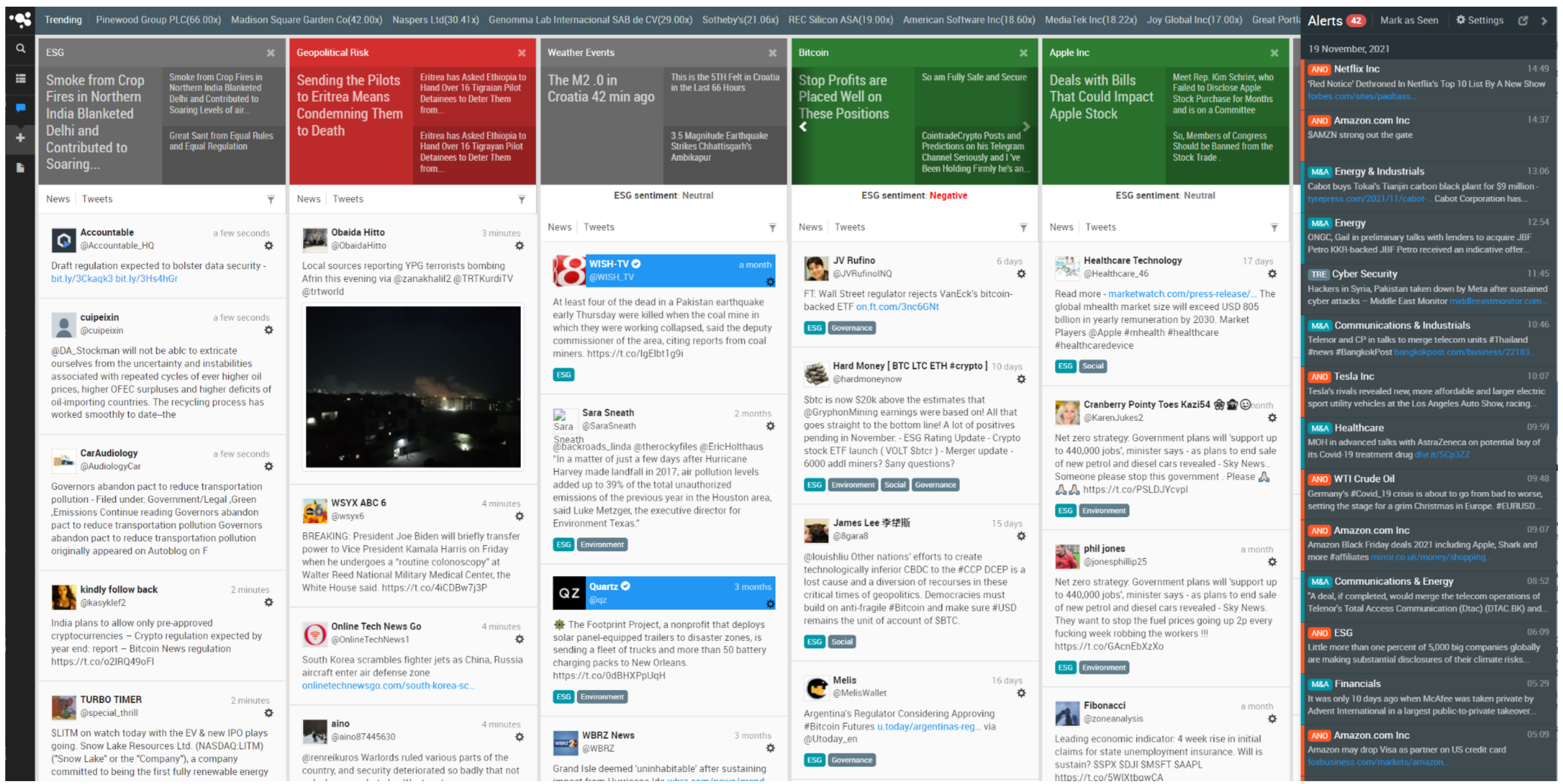
Knowsis offers alternative data solutions and smart real-time alert dashboards

Real-time Data

Competitive pricing

Flexible Use-cases

Access up to 10 years of intraday sentiment scorings for 4,500+ assets, extracted from a variety of online mediums such as the news and Twitter



Query Builder Tool: Customize Knowsis time-series data using query builder to uncover unique predictive insights. Bookmark custom data queries to user's profile for continued monitoring

Alerts: Real-time notification of important changes in Sentiment Scorings

Screener: Use a combination of Knowsis metrics to track fast-growing companies and discover new investment opportunities

Compatibility: Access Knowsis data programmatically via our dynamic API, or export datasets to CSV from the web interface

Offering and Data Access

Sentiment, News and Alerts Query API

- Historical data
- From maximum granularity (single tweet sentiment) to customisable frequency

Sentiment, News and Alerts Stream API

- Real-time data via websocket
- Maximum level of granularity

Sentiment, News and Alerts CSV File

- Full length historical from 2015 (longer upon request)
- From maximum granularity (single tweet sentiment) to customisable frequency
- Various frequencies upon request

USE CASES

- Internal due diligence platform
- Trading Strategies Back testing and factor integration

- Real-time alpha strategies integration
- Real-time alerts and signals
- Alternative news stream and market alerts

- Model factor integration and event study
- Large scale strategy back testing

API Libraries

- **Authorization via Licence token**

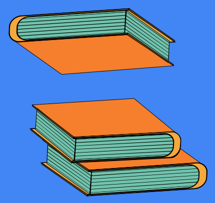
- **Real-time ESG**
  - The WebSocket feed via bidirectional protocol
  - Subscription to ESGStack Dashboard

- **Historical ESG**
  - CSV File
  - HTTP protocol
  - API Query

API Documentation

Link HERE

What's coming next?



Textual Data sources

- Reddit Data
- Reviews Data
- Global News Providers
- Proprietary data sets

Alternative Analytics

- Sentiment Indexes
- Social Volume-Adjusted metrics
- Risk Management analytics
- Smart alerts
- Forecasting models

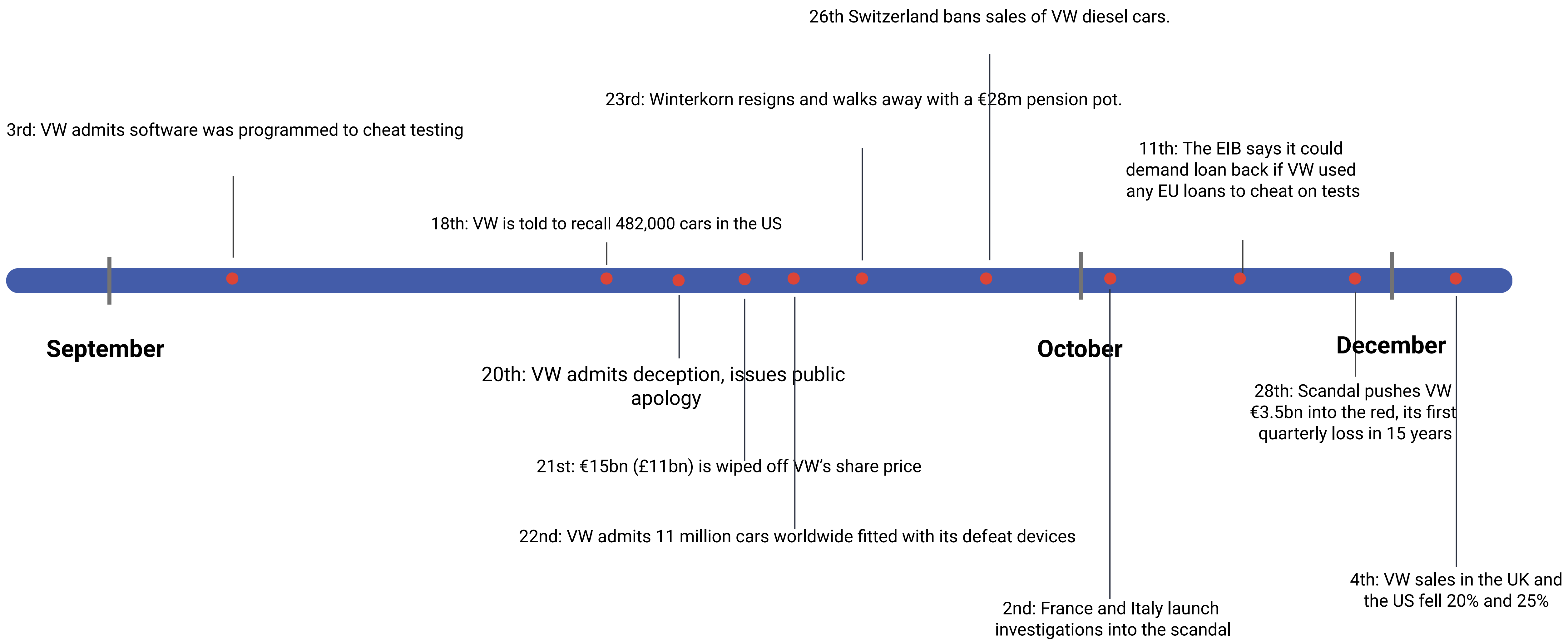
Product Development

- Crypto sentiments
- Crypto new projects discovery model
- Macro Watch (Inflation, Business Cycles, Central Banks)
- Meme-Stocks Monitor and risk-management
- Political Risk and Policy Uncertainty

ESG Case study

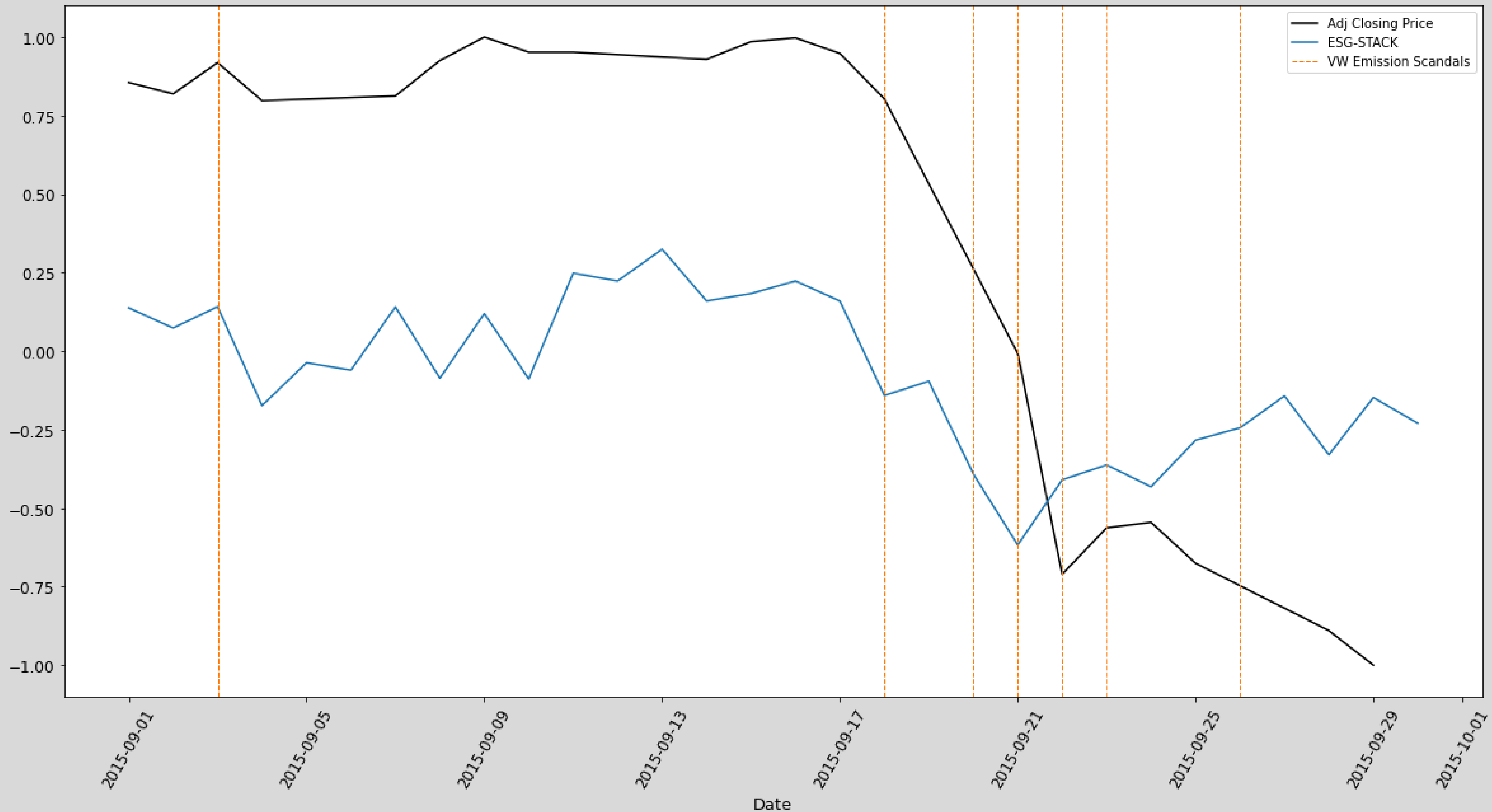
Volkswagen AG (VOW3.DE)

Volkswagen Emissions Scandal (Sept 2015 - December 2015)



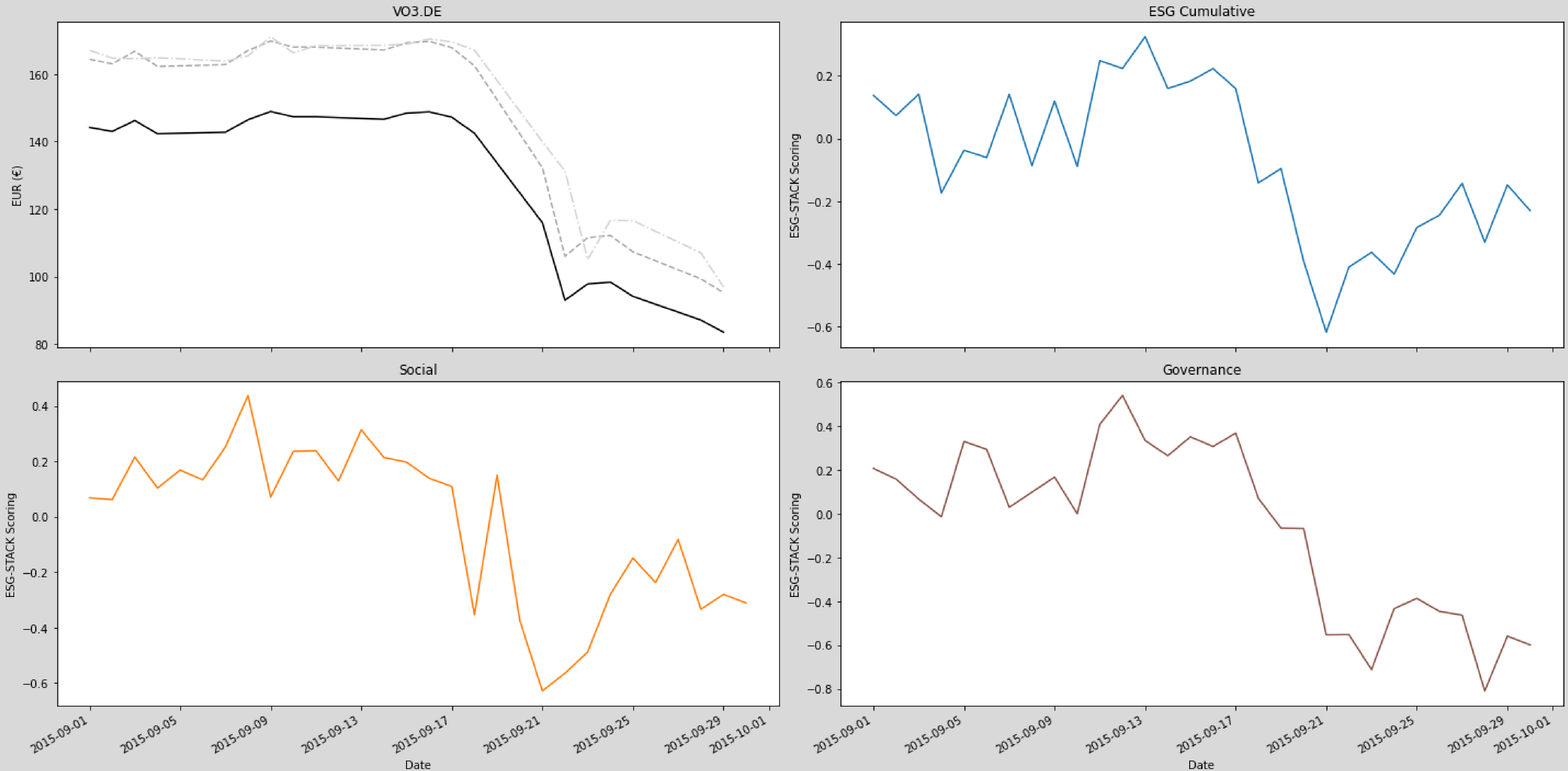
For a discussion of these results please look at Section 3.1 (pp.14-18) of the White Paper: “Modelling ESG with Twitter Data”.

Stock Price and ESG-Cumulative Scorings (VolksWagen AG, Sept 2015)



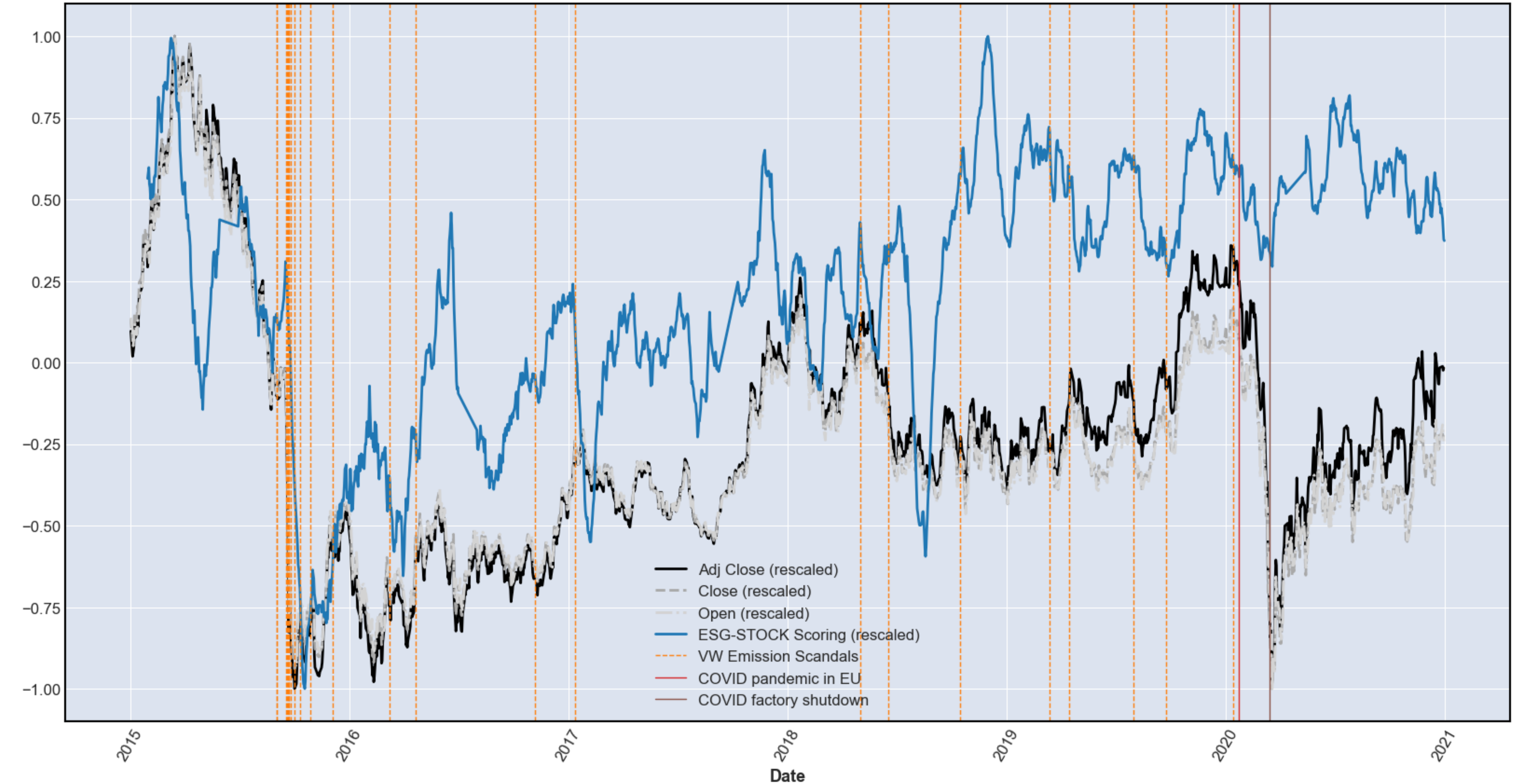
For a discussion of these results please look at Section 3.1 (pp.14-18) of the White Paper: "Modelling ESG with Twitter Data".

Stock Price and ESG-Cumulative, S, and G Scorings (VolksWagen AG, Sept 2015)



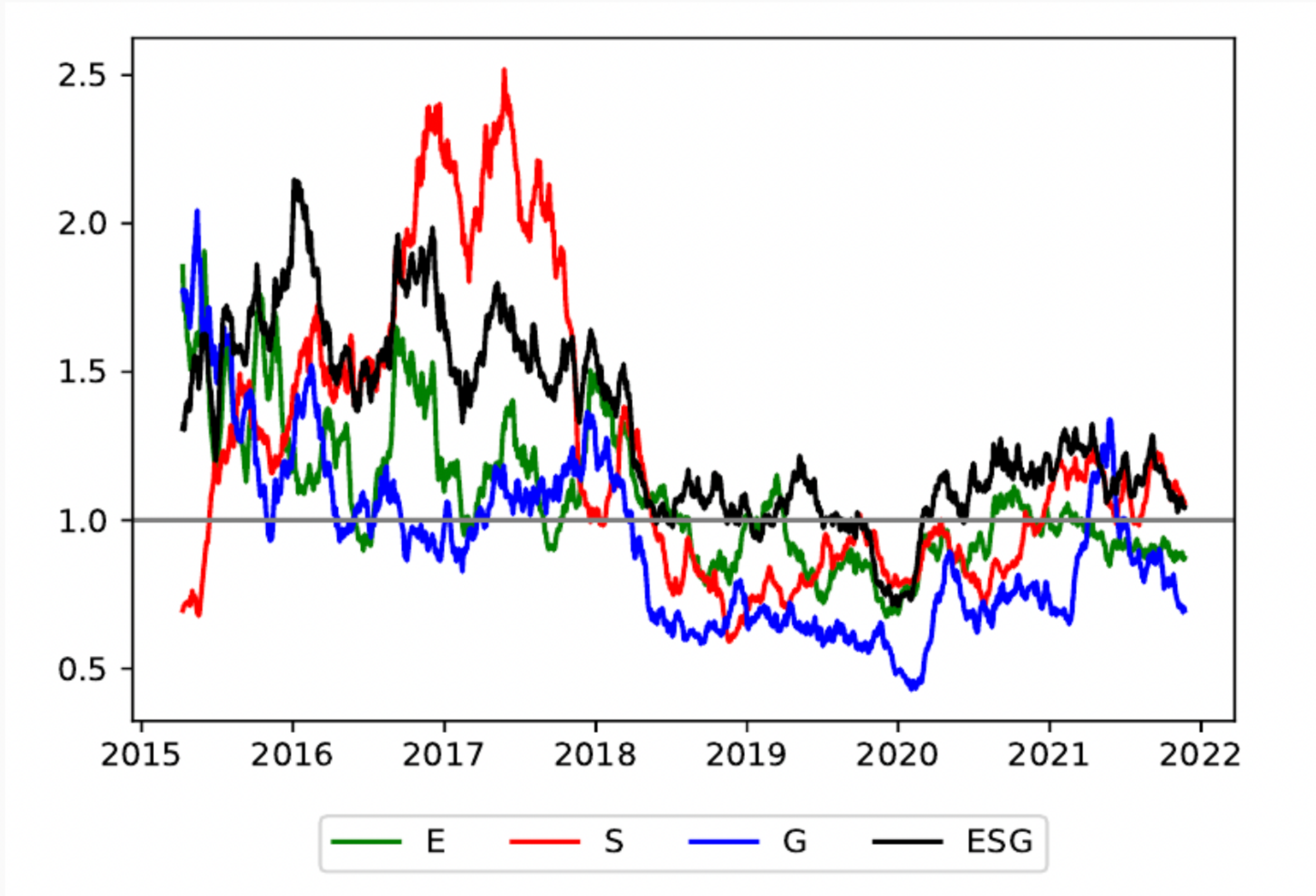
For a discussion of these results please look at Section 3.1 (pp.14-18) of the White Paper: "Modelling ESG with Twitter Data".

Volkswagen: 2015-2020



For a discussion of these results please look at Section 3.1 (pp.14-18) of the White Paper: “Modelling ESG with Twitter Data”.

Ratio between Tesla's Scores and Volkswagen's Ones



We compute the ratio between Tesla's score for E, S, G and ESG and their corresponding scores for Volkswagen. According to our metrics, Tesla enjoyed a significant advantage over Volkswagen, especially on **environmental and social** factors, from 2015 to 2018.

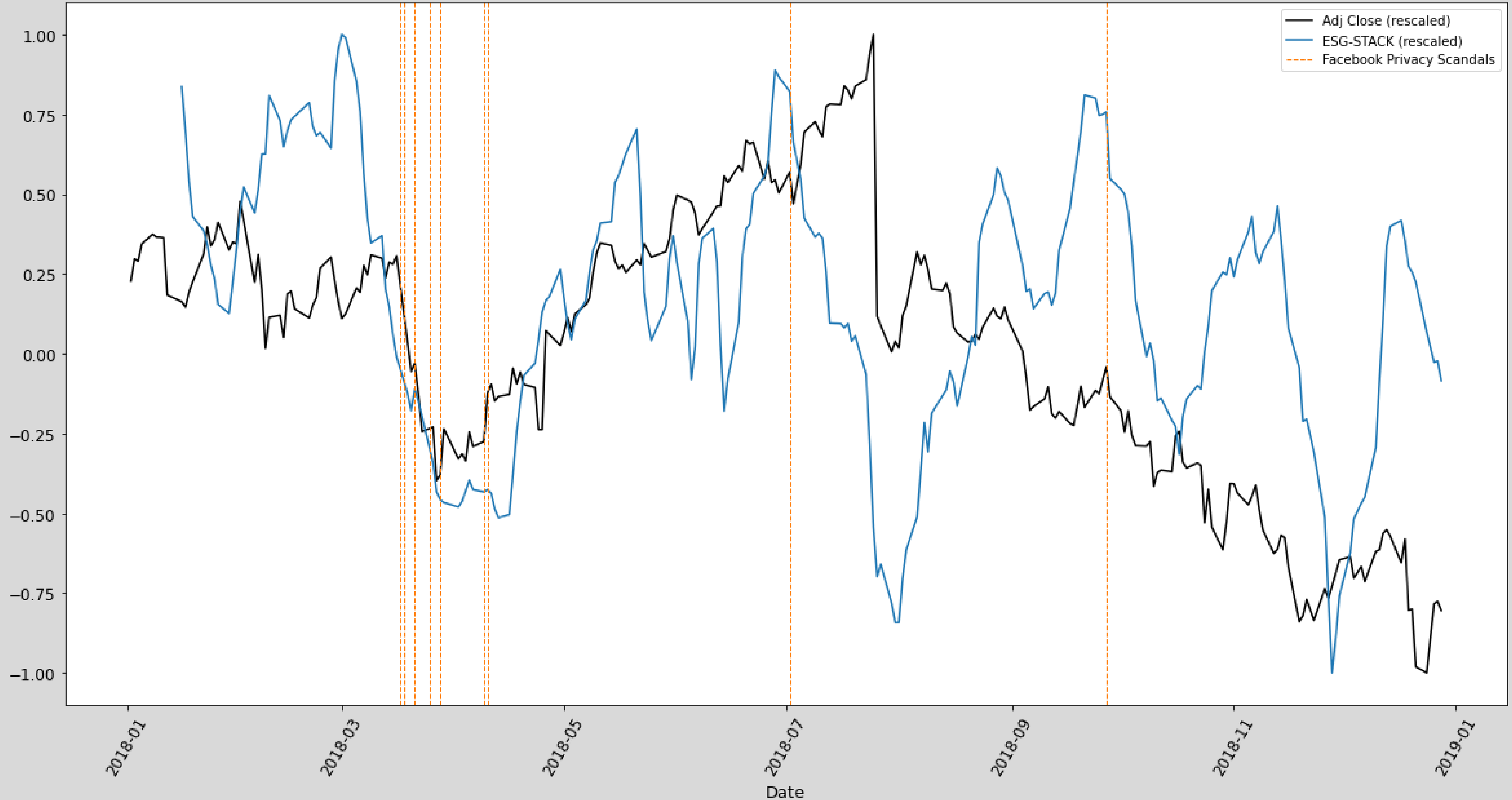
However, from 2018 onward, Volkswagen appears to have closed the gap with Tesla on any ESG dimension.

For a discussion of these results please look at Section 3.1 (pp.14-18) of the White Paper: "Modelling ESG with Twitter Data".

ESG Case study

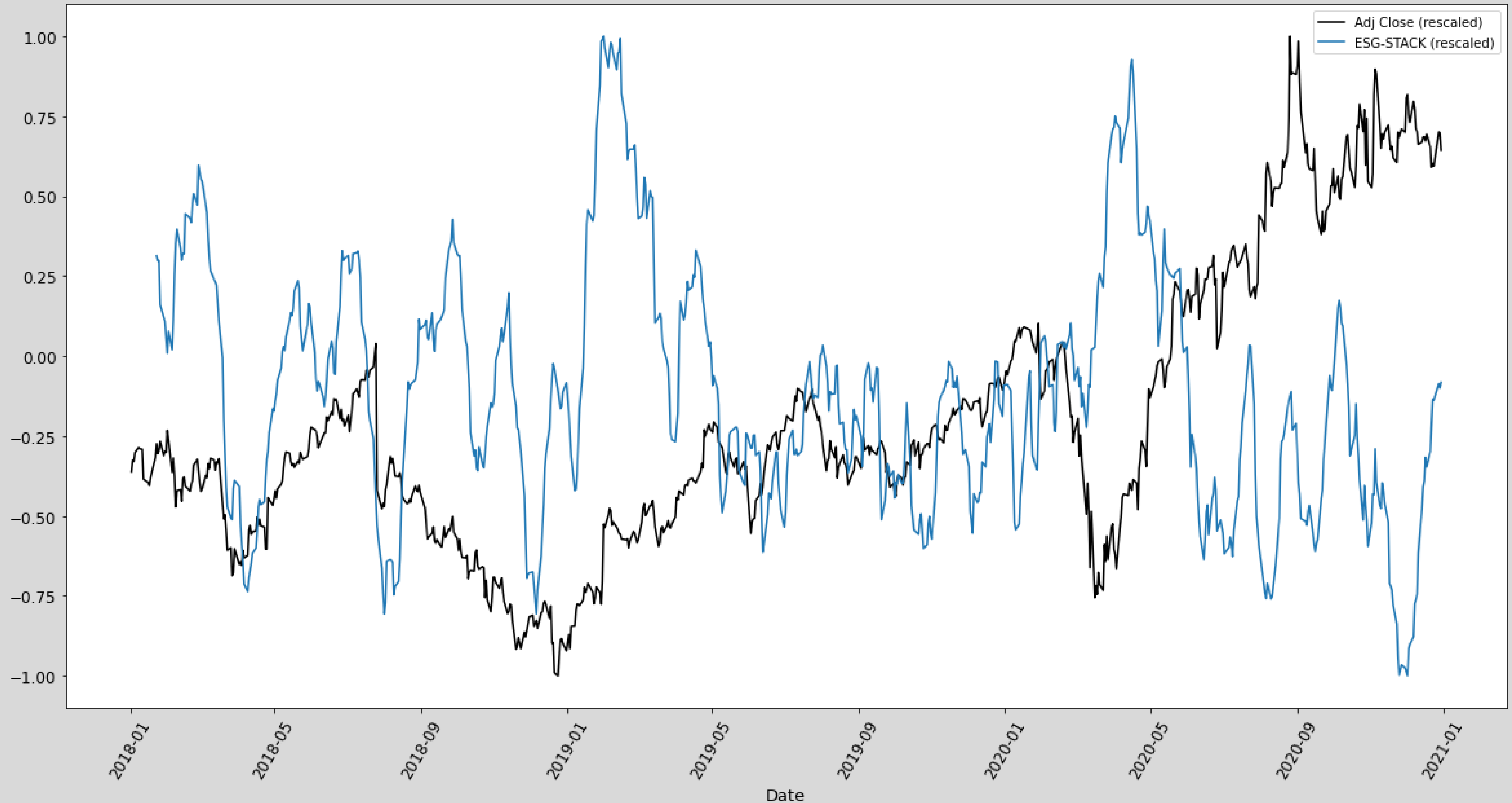
Facebook
(FB:NASDAQ)

Facebook Privacy Scandal (2018)



For a discussion of these results please look at Section 3.1 (pp.18-21) of the White Paper: “Modelling ESG with Twitter Data”.

Facebook (2018-20)



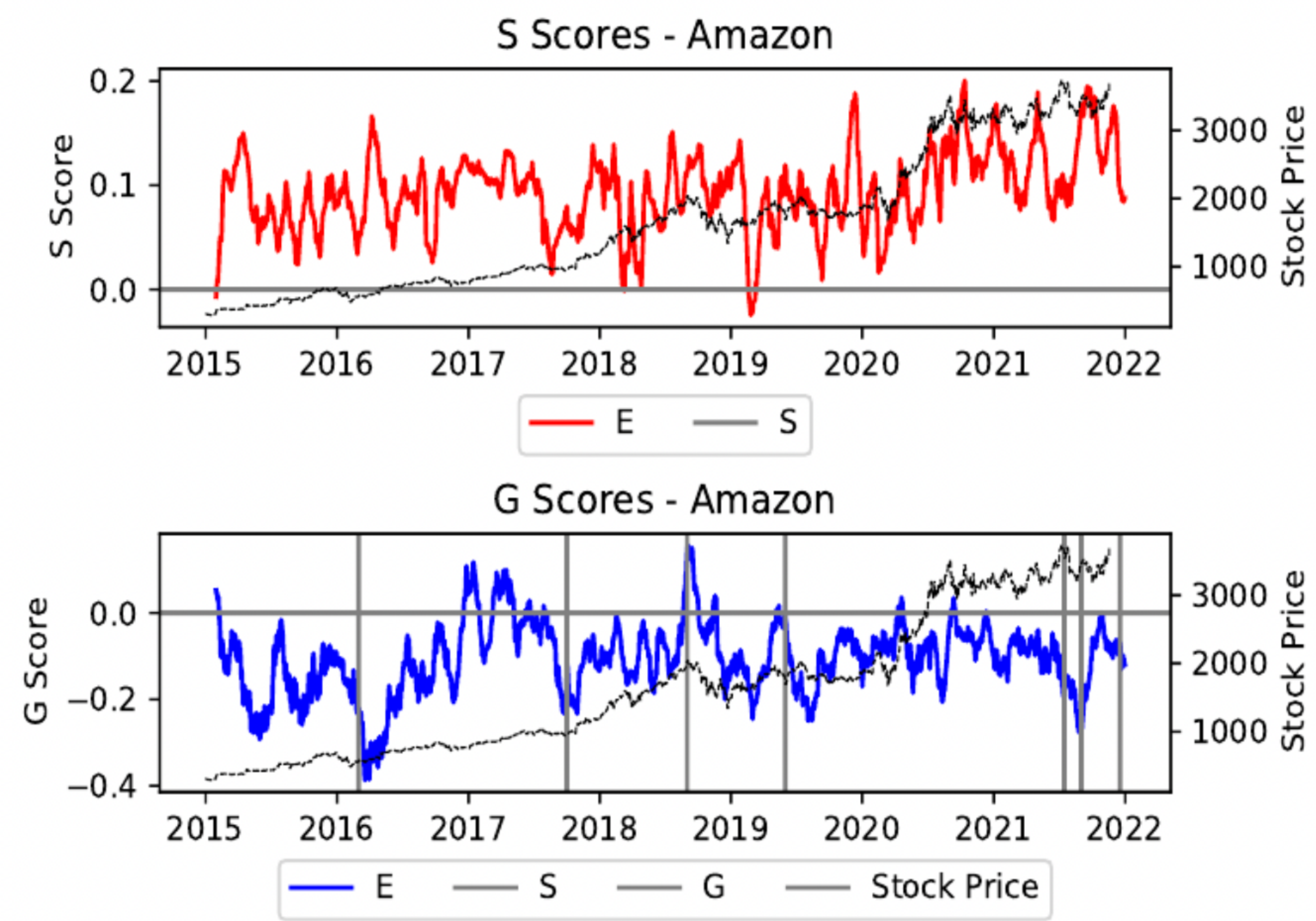
For a discussion of these results please look at Section 3.1 (pp.18-21) of the White Paper: “Modelling ESG with Twitter Data”.

ESG Case study

Amazon

(AMZN:NASDAQ)

An Overview of Governance and Social Scores

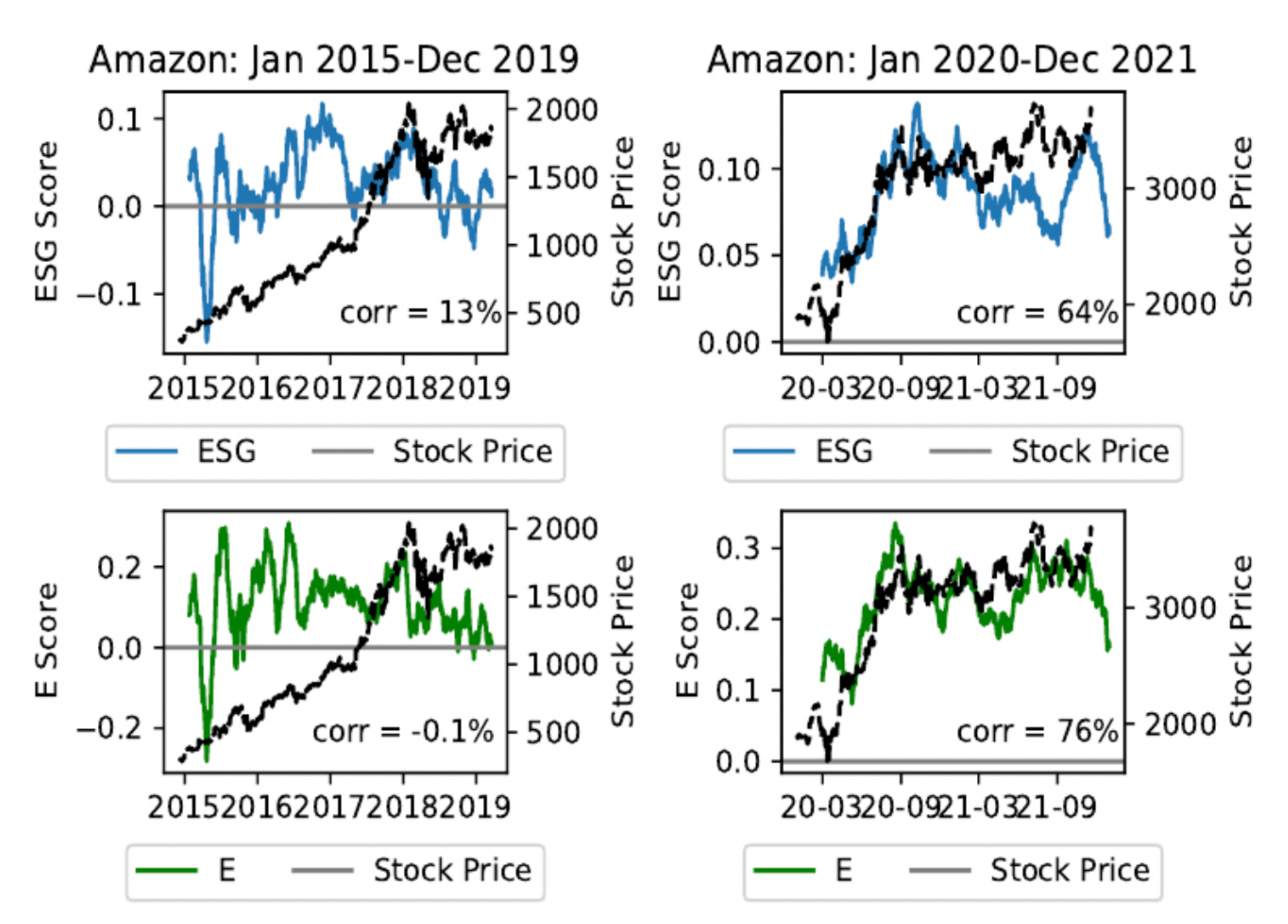


Although Amazon has been frequently accused of neglecting workers' needs, our S-score seems to "promote" Amazon's social policy (e.g. AmazonSmile.

Unlike "S", "G" shows a persistent downward bias. This downtrend is primarily driven by contentious corporate governance events concerning tax-related issues and antitrust cases. We refer to the paper for further analysis.

For a discussion of these results please look at Section 3.1 (pp.21-23) of the White Paper: "Modelling ESG with Twitter Data".

ESG and E-scores in Subperiods of Time



Amazon's environmental performance. First, while the ESG and E scores are little correlated with the stock price from 2015 to 2019, the correlation rockets in the sub-period 2020-2021.

In our opinion, this is preliminary evidence that investors are increasingly embedding ESG considerations into stock valuations and investment decisions.

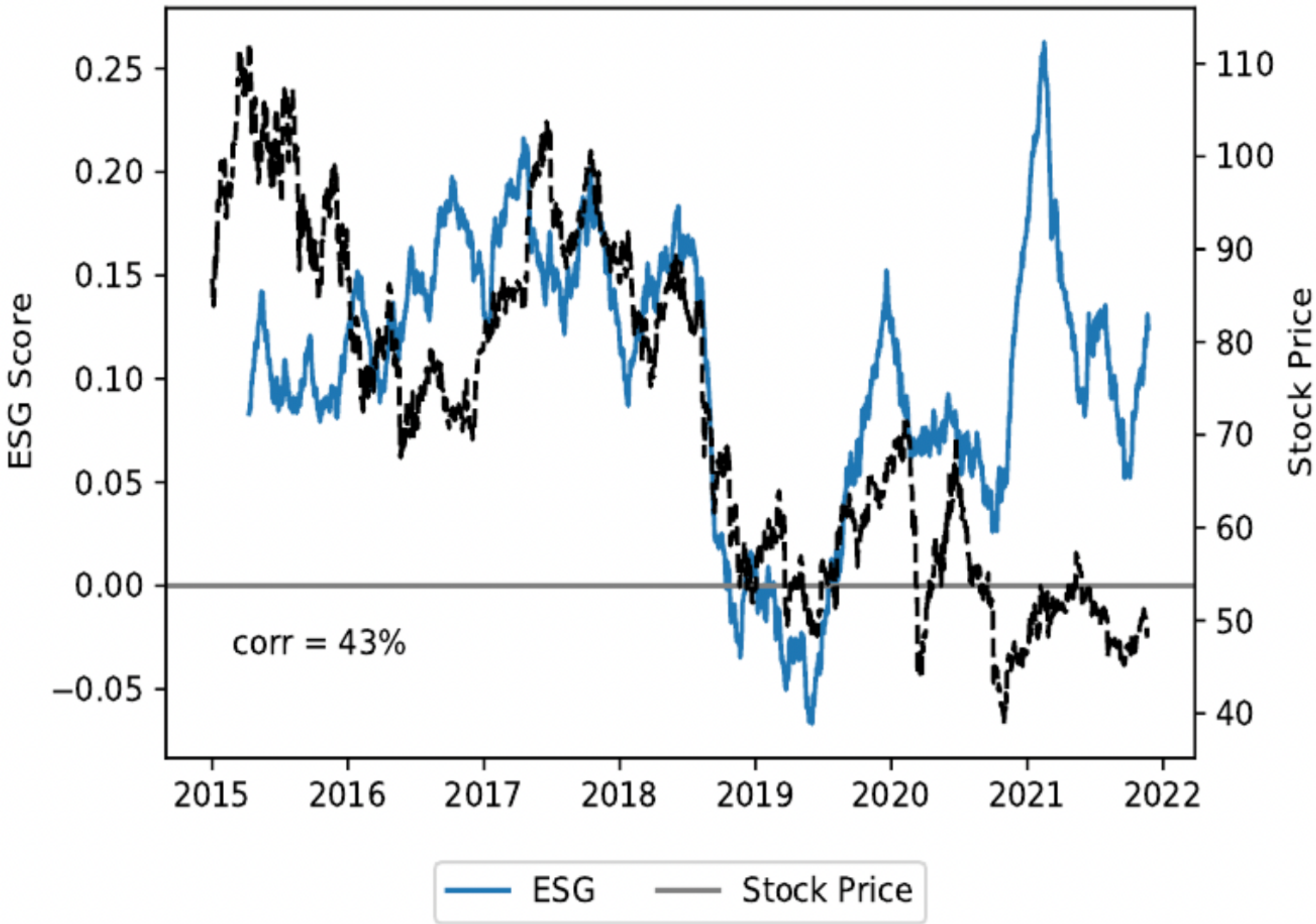
For a discussion of these results please look at Section 3.1 (pp.21-23) of the White Paper: "Modelling ESG with Twitter Data".

ESG Case study

Bayern AG

(BAYN: ETR)

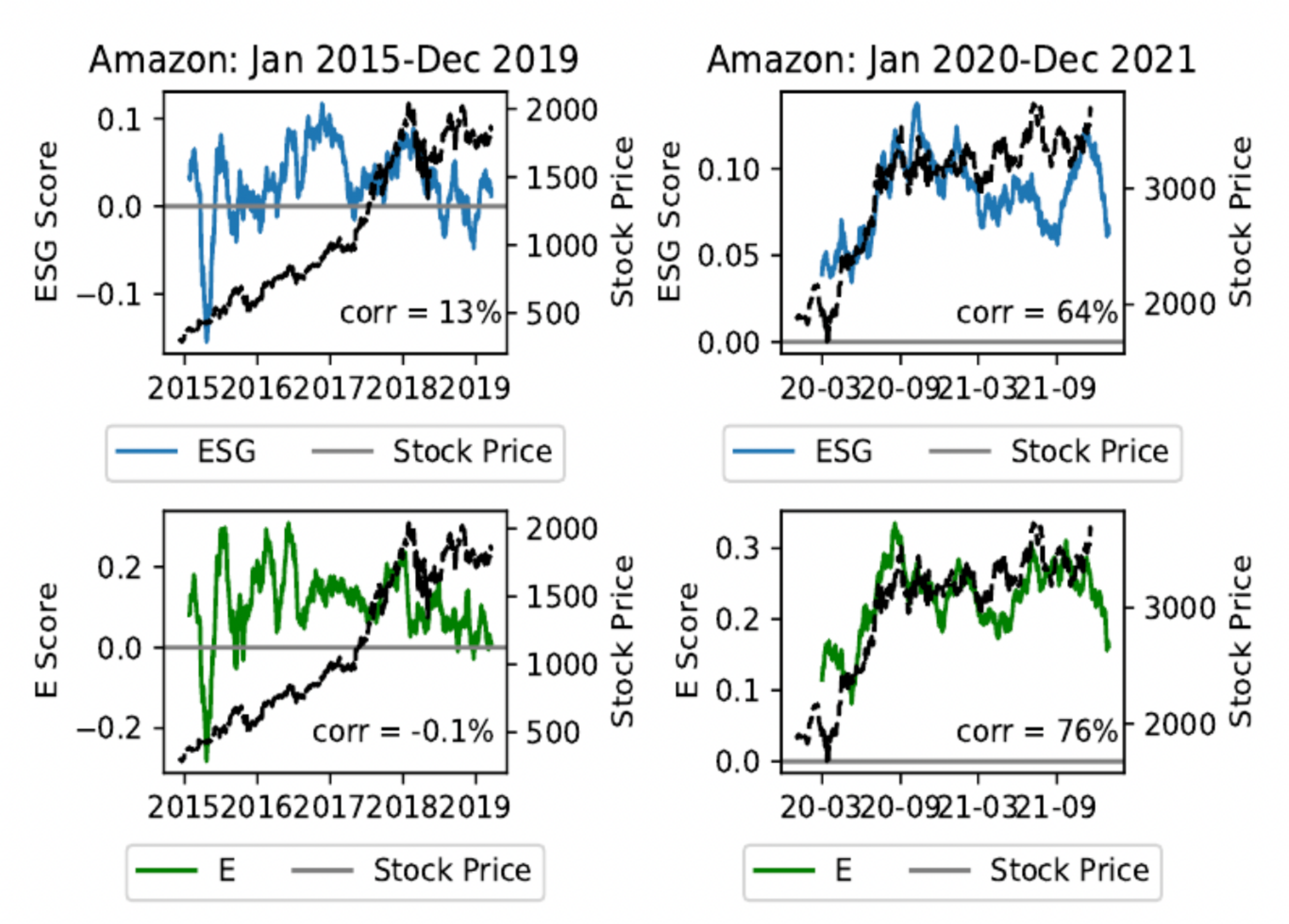
ESG Score vs Stock Price (Full-Sample)



There is a 43% correlation between our ESG score for Bayer and the evolution of its stock price.

For a discussion of these results please look at Section 3.1 (pp.23-26) of the White Paper: "Modelling ESG with Twitter Data".

G-scores at different points in Time

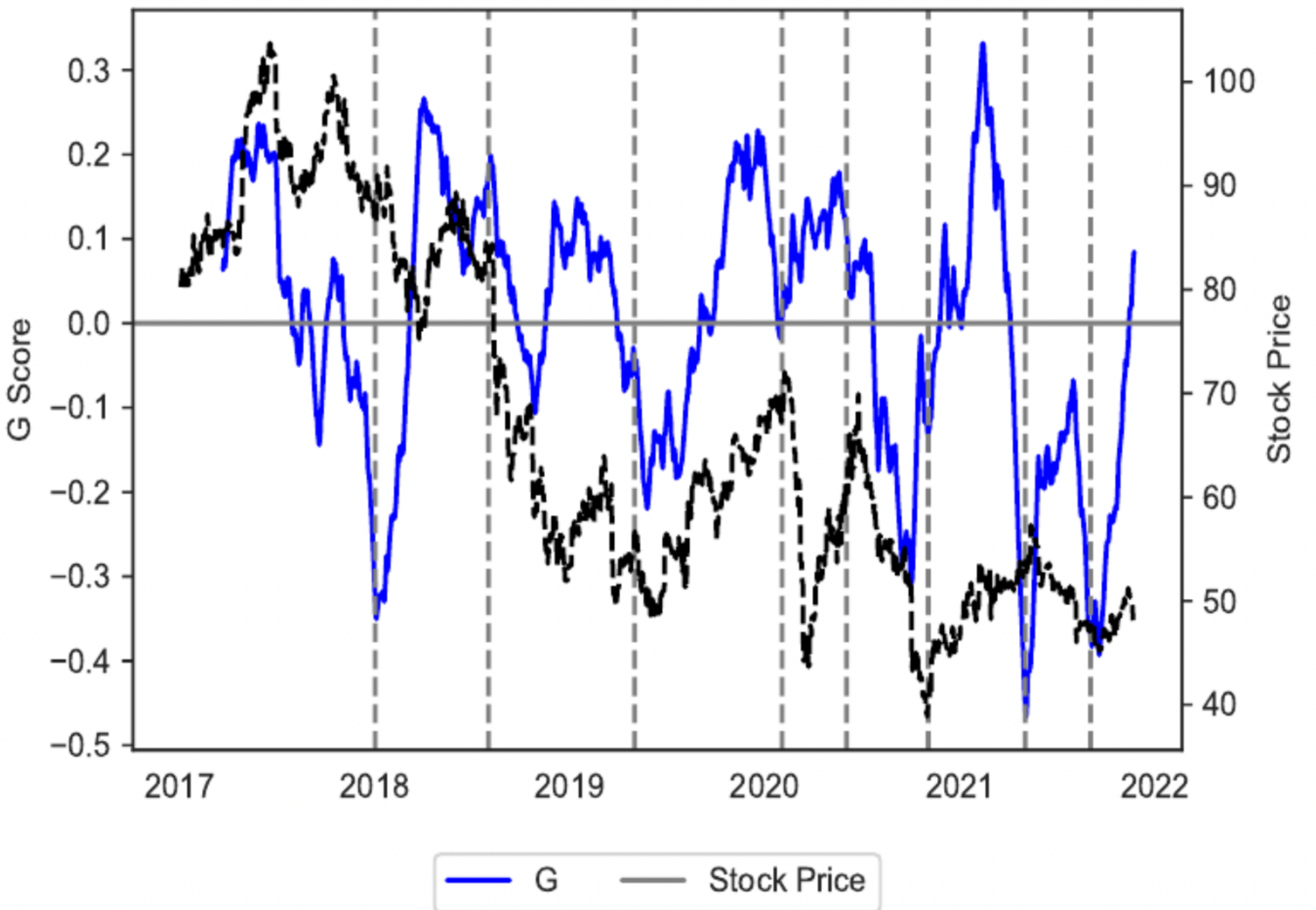


Amazon's environmental performance. First, while the ESG and E scores are little correlated with the stock
price from 2015 to 2019, the correlation rockets in the sub-period 2020-2021.

In our opinion, this is preliminary evidence that investors are increasingly embedding ESG considerations into stock valuations and investment decisions.

For a discussion of these results please look at Section 3.1 (pp.21-23) of the White Paper: "Modelling ESG with Twitter Data".

Stock Price vs G-Score



Since Bayer has been affected by various ESG-relevant episodes, we study whether our model is suited to capture those events. The figure displays the results for the period 2017-2021.

For a discussion of these results please look at Section 3.1 (pp.23-26) of the White Paper: "Modelling ESG with Twitter Data".

Key Takeaways

1) Our metrics are proven to be extremely sensitive to environmental, social and governance events, responding in real-time and often *ahead of the market*.

2) Correlation coefficients are not stable throughout the period time under study for the majority of the stocks. This is however an expected finding since our model isn't calibrated to respond to any generic, despite meaningful, event for the stock price. ESG factors do not always drive price dynamics. Rather than tracking the stock price closely, we believe that our metrics work at their finest as a real-time anomaly detector, capturing ESG relevant momenta.

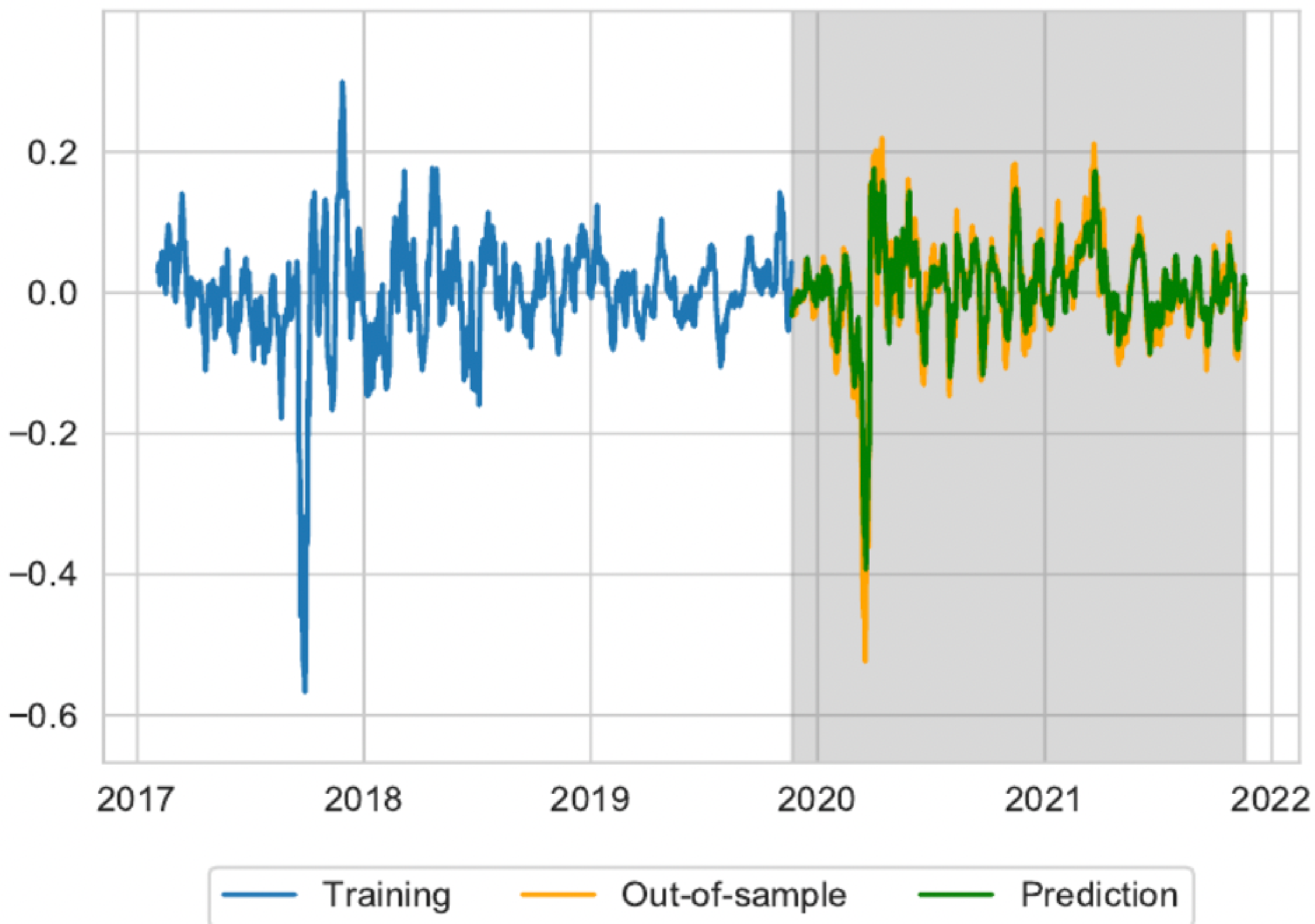
4) Although correlation coefficients might often be low from 2015 until 2021, they are significantly higher when the sample size is narrowed to be within a year of time frame (see, for example, Amazon's example). This suggests that, with a shorter timespan, our score can also shed light on share prices' dynamics.

5) For many stocks we analyzed, correlation coefficients tend to increase over the last couple of years. We believe this is not a fluke but it is rather likely due to the increasing integration of ESG factors into stock valuations and portfolio strategies.

Advanced Applications

Time-series Forecasting

Forecasting Volkswagen's Stock Return

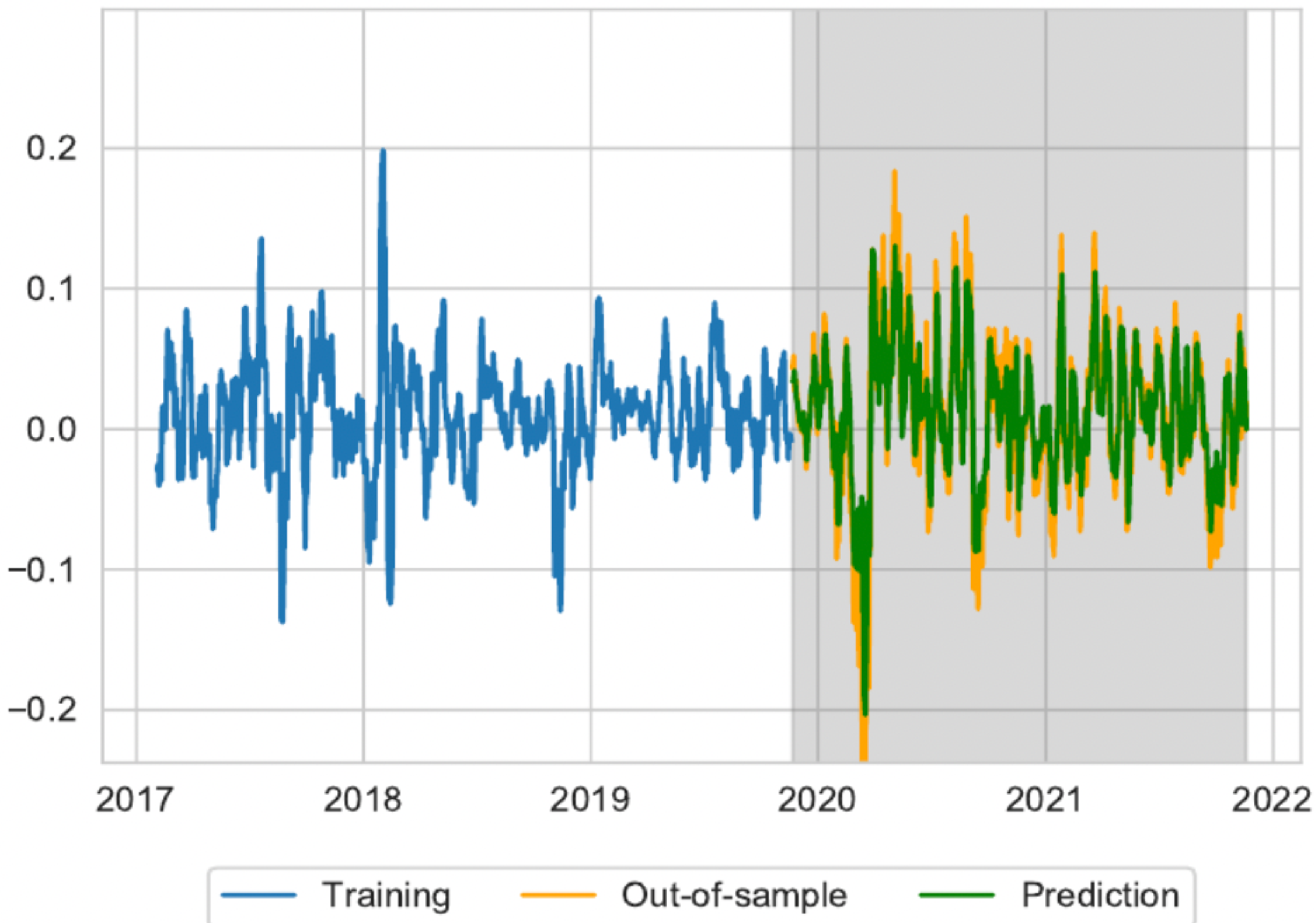


We forecast, using an LSTM model, Volkswagen's stock returns for the period 2020-2021.

The forecast horizon is rather challenging from a forecasting standpoint. In fact, many well-established and long-standing time series models have failed to keep up with the evolution of the pandemic.

Although our ESG metrics can sometimes drift away from the evolution of the stock price, they can still provide a valuable integration to existing forecasting models.

Forecasting Facebook's Stock Return



We forecast, using an LSTM model, Facebook's stock returns for the period 2020-2021.

The forecast horizon is rather challenging from a forecasting standpoint. In fact, many well-established and long-standing time series models have failed to keep up with the evolution of the pandemic.

Although our ESG metrics can sometimes drift away from the evolution of the stock price, they can still provide a valuable integration to existing forecasting models.

Integrationg ESG into the Investment Process

Although it is beyond the scope of this paper to provide a fully-fledged model that integrates our ESG metrics into a comprehensive framework, we want to briefly touch upon a few potentially relevant applications for the ESG investment community.

1) From a qualitative perspective, our model could help improve on “ESG red flags” as well as on “ESG scorecards”. As for the former, our model is equipped to flag emerging ESG risks and threats at a very early stage, enabling investors to capture the speed and intensity of their possible development. Concerning the latter, instead, our model would produce different moving averages as an additional score to be used in absolute and relative terms against sectors, geographic regions and index benchmarks. Those additional metrics would help “scorecards” to have a more continuous computation and relative signalling power.

2) From a quantitative point of view, we notice that the scoring momentum is a potentially valuable additional factor for either factor model portfolios or as a trigger for exclusionary-screening adjustments (see, for instance, PRI). Furthermore, from an active ownership standpoint, the model detects in real-time areas of further discussion and engagement with investee’s companies and alerts them to respond timely to issues raised on social media.

Info and contacts

Learn More

www.knowsis.com
esgstack.com

hello@knows.is

Tommaso Iaquone: Head of Growth
tommaso@knows.is

Filippo Lanza: Director
filippo@knows.is

Follow us

Knowsis