



Let's get physical: Comparing metrics of physical climate risk[☆]

Linda I. Hain, Julian F. Kölbel, Markus Leippold^{*}

Department of Banking and Finance, University of Zurich, Plattenstrasse 14, 8032 Zurich, Switzerland

ARTICLE INFO

JEL classification:

G12
G24
G32
Q54

Keywords:

Climate change
Physical risk scores
Disagreement
Model uncertainty

ABSTRACT

Investors and regulators require reliable estimates of physical climate risks for decision-making. While assessing these risks is challenging, several commercial data providers and academics have started to develop firm-level physical risk scores. We compare six physical risk scores. We find a substantial divergence between these scores, also among those based on similar methodologies. We show how this divergence could cause problems when testing whether financial markets are pricing physical risks. Hence, financial markets may not adequately account for the physical risk exposure of corporations using available risk scores. Finally, we identify key sources of uncertainty for further investigation.

1. Introduction

The physical impacts of climate change pose major economic risks. According to [Swiss Re Institute \(2020\)](#), weather-related events such as storms, floods, droughts, and wildfires caused global annual economic losses of USD 212bn on average over the last decade, and climate change is expected to lead to more frequent and more severe extreme weather events in the future.¹ Consequently, financial institutions aim to integrate these additional risks in financial decision-making. To do so, investors depend on reliable information about a firm's exposure to this risk. However, companies' disclosure of climate-related risks so far is limited and voluntary in most countries. [TCFD \(2019\)](#) and [SASB \(2018\)](#) offer guidance in climate risk disclosure. Nevertheless, the standardization of climate risk disclosure is still evolving. Moreover, the lack of information about climate risks remains a critical constraint to reflect the risks of climate change in financial decision-making ([Behnam et al., 2020](#)).

In response, several commercial providers have begun to develop climate risk scores to provide investors with firm-level indicators of physical climate risk. Some of the large credit-rating agencies entered the race to provide climate-risk scores, representing a significant step towards financial institutions committing to the idea that climate change should be a routine consideration when evaluating any company's financial strength.² Therefore, climate risk scores have the potential to steer capital towards climate-resilient business models and incentivize corporations to pursue resiliency.

However, [Fiedler et al. \(2021\)](#) argue that the scientific base for climate risk scores is still developing, raising the question of whether the use of these scores sufficiently improves financial decision-making. To explore this question, we compare six scores

[☆] We thank Samuel Vigne (the editor), Julia Bingler, Jeffrey Bohn, Jim Hawley, Florian Heeb, Jonathan Krakow, Nico Kröner, Violaine Lepousez, Liall Medina, Falko Paetzold, Oliver Schelske, and the participants at the UNEP FI Climate metrics for physical risks seminar, the Swiss Finance Institute (SFI) research days, the University of Zurich, and the Swiss Re Institute for helpful comments and inspiring discussions. The financial support of the Swiss Re Institute is gratefully acknowledged. Julian also gratefully acknowledges financial support of the BMW Foundation Herbert Quandt.

^{*} Corresponding author.

E-mail addresses: lindaisabella.hain@bf.uzh.ch (L.I. Hain), julian.koelbel@bf.uzh.ch (J.F. Kölbel), markus.leippold@bf.uzh.ch (M. Leippold).

¹ See, e.g., [IPCC \(2014\)](#), [Stott \(2016\)](#), [Herring et al. \(2015\)](#), [Sun et al. \(2014\)](#).

² See, <https://www.ft.com/content/59f60306-d671-11e9-8367-807ebd53ab77>.

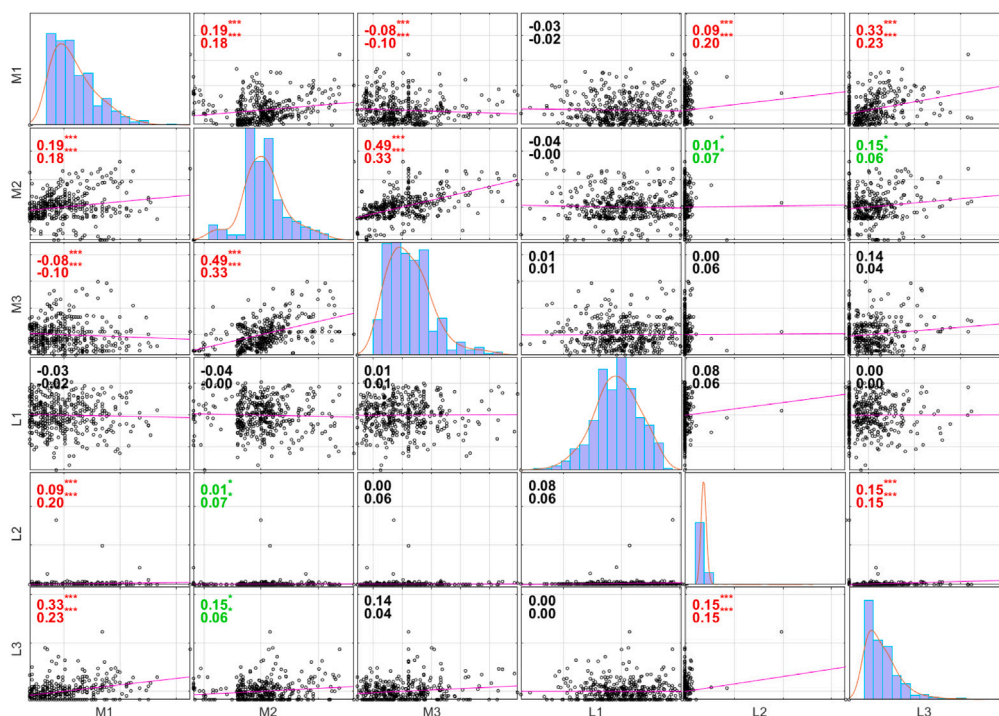


Fig. 1. Correlation measures and Distributions of scores: Horizon 2050, RCP8.5. This graph shows the distribution of each score in the diagonal panels. In the off-diagonal panels it shows scatterplots, Pearson correlation coefficient (first row), and Kendall's tau (second row) for each score pair. Model-based scores reflect the RCP8.5 climate change scenario. (** $p < 0.01$ in red, ** $p < 0.05$ in blue, * $p < 0.1$ in green). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

for firm-level physical climate risks to determine to what extent scores from different providers converge. The metrics focus on the physical risks of climate change, defined as a firm's additional exposure to hazards associated with a changing climate. Hazards include acute hazards such as floods, droughts, storms and chronic hazards such as sea-level rise.

2. Data

In our analysis, we include model-based approaches that model corporations' risks on the basis of climate models and language-based approaches that quantify risk on the basis of textual sources about corporations' climate risk exposure. Model-based approaches include the commercial scores from *Trucost* (M1), *Carbon 4 Finance* (M2) and *Southpole* (M3). Language-based approaches include a commercial score from *Truvalue Labs* (L1), and two academic scores, *Firm-level Climate Change Exposure* (L2) (Sautner et al., 2020a) and *BERT-based climate risk measure* (L3) (Kölbel et al., 2020). L1 relies on news articles as a text source, L2 on earnings call's transcripts, and L3 on regulatory filings (see Table A.1 in the Appendix for details on the methodologies).

Our sample consists of 408 US corporations included in the *S&P 500*, the *Dow Jones Industrial Average* (DJ30) and the *Dow Jones Composite Average* (DJ65).³ We ensure that scores are comparable with regards to objective and assumptions. All scores represent the available information at the end of 2019. For model-based scores, we compare scores that aggregate different climate hazards and assess physical risk under an RCP8.5 scenario for a time horizon until 2050.⁴ Language-based scores were averaged over time to facilitate comparison (Table A.1).

3. Score correlations

Fig. 1 displays the distributions of climate risk scores in the diagonal panels and correlations in the off-diagonal panels. Pearson correlations range between -0.10 and 0.49 (Kendall's tau -0.08 to 0.33), indicating overall a low level of agreement between scores over the sample. Six out of 15 correlations are significantly different from zero at the 1%-significance level for Pearson's correlation and Kendall's tau. The overall divergence also shows up visually in the scatterplots.

³ In total, we have 511 firms and 408 of them are covered by all climate-score providers.

⁴ All our results remain robust when we analyze the RCP4.5 scenario. These results can be obtained from the authors.

Table 1
Sector ranking: Horizon 2050, RCP8.5.

Sector	M1	M2	M3	L1	L2	L3	Avg	SD
Health Care	9	6	7	10	9	11	8.7	1.9
Consumer Staples	7	5	5	4	8	3	5.3	1.9
Consumer Discretionary	3	8	9	6	6	6	6.3	2.1
Information Technology	10	7	6	5	10	9	7.8	2.1
Energy	4	2	1	7	5	4	3.8	2.1
Materials	5	3	2	8	3	2	3.8	2.3
Industrials	6	4	8	2	2	7	4.8	2.6
Communication Services	8	10	10	3	11	10	8.7	2.9
Financials	11	9	3	1	7	8	6.5	3.8
Real Estate	2	11	11	9	4	5	7.0	3.8
Utilities	1	1	4	11	1	1	3.2	4.0

This table shows sector rankings (11 sectors) for each score, average of score ranks (Avg), and standard deviation (SD) across score rankings for each sector. The sector ranks are based on the average score value per sector.

Within model-based scores, **M2** is significantly and positively correlated with **M1** and **M3**, but with a very low correlation. The correlation between **M1** and **M3**, however, is significantly negative, implying not just a divergent but slightly opposite view on climate risk. Within the language-based scores, the correlation between **L2** and **L3** is significant but weak, whereas **L1** does not significantly correlate with any other score. These observations change only marginally when considering an RCP4.5 scenario.⁵

Potentially, different providers arrive at a similar ranking within sectors. We therefore analyze the rank correlations within sectors as defined by the Global Industry Classification Standard (GICS) and present the results in [Appendix B](#). Overall, within-sector correlations also indicate a substantial divergence of physical risk scores. Hence, no sector emerges for which the providers exhibit more agreement on physical climate risk exposures, also not for *Real Estate*, although the literature argues that physical risks in real estate are comparatively easier to model and better understood ([Baldauf et al., 2020](#); [McNamara and Keeler, 2013](#)).

4. Sector rankings

While correlations are low within sectors, scores might agree on which parts of the economy are most affected by physical climate risk. [Table 1](#) shows that, in terms of average ranking, *Utilities*, *Energy*, and *Materials* are considered as the most exposed sectors.⁶ This finding seems reasonable, given that all these sectors rely heavily on fixed tangible assets exposed to natural hazards. *Health Care*, *Communication Services* and *Information Technology* turn out to be the least exposed to physical climate risk.

Calculating the standard deviation (SD) across the six score rankings for each sector and interpreting SD as a measure of model uncertainty, we find substantial differences in uncertainty across sectors. *Health Care*, *Consumer Staples* and *Consumer Discretionary* have the lowest dispersion across providers. In contrast, *Financials*, *Utilities*, and *Real Estate* exhibit a large degree of dispersion. *Real Estate* has the second-highest standard deviation which is in line with the low within-sector correlations for this sector (see [Table B.1](#)). For *Utilities*, the six scores display the greatest standard deviation, which is driven by **L1**. While four of the six providers agree that this sector is ranked as the most exposed to physical risk, **L1** shows exactly the opposite ranking.

5. Portfolio sorts

A common approach in the asset pricing literature to investigate whether the financial market indeed prices a risk factor is sorting stocks into high- and low-risk portfolios, which also has been applied by [Bansal et al. \(2016\)](#) and [Bolton and Kacperczyk \(2020\)](#) to climate change risks recently. To illustrate the potential consequences of the significant disagreement in risk assessments, we examine the robustness of the sorting mechanism to the use of different scoring models. To this end, we draw Venn diagrams, which visualize the possible logical relations between different models and approaches.

[Fig. 2](#) presents the different Venn diagrams for which we focus on the (commonly used) upper quintile and tercile. With complete agreement, the Venn diagrams would collapse into a single circle. However, we find that there is very little agreement both within and across approaches. For instance, for the model-based approaches under the RCP 8.5 (RCP 4.5) and using quintiles, we find only 10 (7) companies that are sorted consistently by **M1** to **M3** into the upper quintile. For a total of 408 companies, the quintile should theoretically contain 82 companies. Hence, we can interpret these numbers as an agreement of 12.2% (8.5%) on the quintile constituents. Switching to terciles, the relative level of agreement for the model-based approaches improves slightly to 14.8% (16.4%).

For language-based approaches, the agreement is even less. For the upper quintile portfolio, **L1** to **L3** only agree on three companies, and for the upper tercile only on ten companies. When comparing the model-based and language-based approaches

⁵ The corresponding analysis can be obtained from the authors.

⁶ We recall that a low score reflects a high risk exposure.

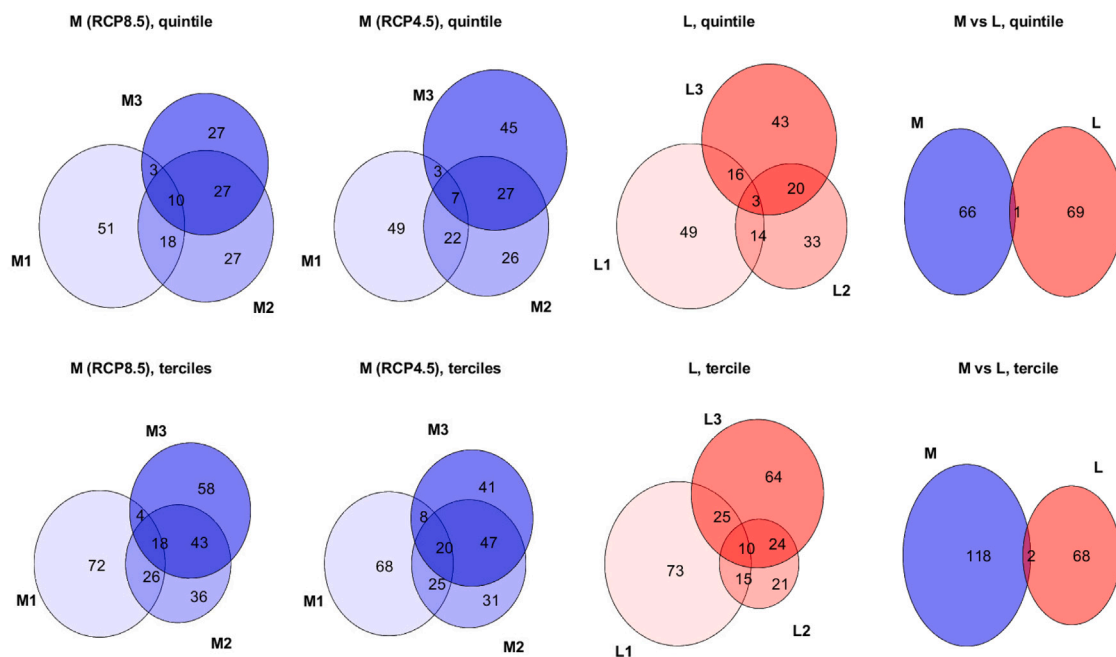


Fig. 2. Venn diagrams for portfolio sort agreement. This graph shows the Venn diagrams for the upper quintile and tercile portfolios. The sample consists of 408 companies in total. Hence, quintiles and terciles ideally consist of 82 and 122 firms. The first three diagrams in each row provide a separate analysis for the model-based and language-based approaches. We compare the model-based approaches **M1** to **M3** for the RCP 8.5 and RCP 4.5 scenarios and the language-based approaches **L1** to **L3** separately. In the last column, we compare the model-based and the language-based models together where the number of companies for model-based scores and for language-based for the respective percentile are displayed and the overall intersection across all six scores for RCP8.5. With perfect agreement, the circles for each Venn diagram would collapse into a single circle. Note that **L3** provides non-zero scores for 70 companies.

jointly, we find only one company that consistently ranks in the top quintile for both approaches, i.e., across all six scores. For the upper tercile, there are two companies consistently ranked. Since the tercile should theoretically contain 122 companies, the agreement is less than 2%. These results show that all inferences drawn from a risk premium analysis for physical climate risk based on portfolio sorts are highly sensitive to the choice of the risk score.

6. Sources of uncertainty

The documented divergence of scores suggests that there remains substantial uncertainty about firms' exposure to physical climate risk. Such a kind of model uncertainty discussed in [Brock and Hansen \(2018\)](#) is to be expected, given the inherent complexity of assessing a firm's physical risk exposure. To reduce this uncertainty going forward, we qualitatively identify the key sources of uncertainty and suggest measures to address them.

We first discuss model-based approaches. All three approaches rely, according to their methodology description, on a basic framework of hazards, exposure, and vulnerability. The first component, hazard, is typically derived from climate models that are publicly available. It is commendable that assessments rely on widely available climate change scenarios. However, [Fiedler et al. \(2021\)](#) point out that there is a mis-match between the coverage of climate models and financial risk analysis requirements. This includes that climate models themselves diverge in their predictions, especially at high spatial resolution and short time frames, which are of particular interest to investors. Thus, transparency about the climate model behind the scores and its limitations is a critical first step.

The second element, exposure, is based on a firm's spatial distribution of physical assets and supply chains. These data are challenging and expensive to collect, and score providers need to resort to proprietary data. Differences in these asset-level data sets could be driving risk assessments apart. For example, the strikingly diverging assessments for the real estate sector could be driven by inconsistent information about asset locations. In order to explore this possibility, researchers would benefit from an open-source data set of asset locations, that could be used to benchmark different methodologies. Potentially, the National Establishment Time Series (NETS) data already applied in [Addoum et al. \(2020\)](#) could serve in this role. Commercial providers of risk scores could use the reference data to benchmark their own methodologies and extend the scientifically tested methods to other data sets that may be larger and proprietary.

The third element, vulnerability at the entity level, is perhaps the least well understood and, therefore, a major source of uncertainty. The same firm may suffer vastly different financial consequences, depending on whether it was prepared for potential exposure. Such a measure of preparedness could include management's awareness of risks, the build-up of response capacity, and

insurance against the risks. It may even be the case that firms emerge more profitable after certain events as shown in Horwitz (2009). A way forward would be to compare scores specifically aiming to measure vulnerability, given that such an assessment is likely to contain a high degree of judgment.

The construction of language-based scores relies crucially on the accessed text sources and the natural language processing (NLP) algorithm that is used to analyze it. We identify three sources of uncertainties for language-based approaches that are likely to drive score divergence. First, language-based scores are limited to the knowledge contained in the analyzed text sources and depend on the applied NLP-algorithm. L1 relies on media sources, L2 on earnings calls transcripts, and L3 on regulatory filings. It is therefore obvious that these scores represent different perspectives and information sets. Another critical point is the type and precision of the algorithms used in the text analysis. Varini et al. (2020) show that keyword-based methods tend to perform worse than contextual algorithms in the task of identifying sentences that address the topic of climate change. Thus, uncertainty in language based scores comes from the text source and the algorithm.

The second source of uncertainty for language-based scores is that they tend to measure a relative rather than an absolute view of physical climate risk. For example, L2 reports the fraction of sentences in an earnings call that were dedicated to physical climate risks (and L3 adopts a similar approach). Other prominent issues, such as COVID19, can crowd out physical climate risk. In other words, the physical risk score is relative to other concurrent issues of concern. A possible solution may be to disentangle climate risks from other risks that threaten the enterprise at the same time. Furthermore, language-based approaches may pick up the relative change over time, given that information that was given previously is not necessarily repeated. These are important ways in which language-based approaches can provide very different scores compared to model-based scores, which attempt to model absolute exposure.

The third source of uncertainty for language-based scores is the unspecified future time horizon and climate change scenario. Currently, algorithms do not detect whether statements refer to a specific scenario or time horizon, and this may in fact not be specified in the text itself. As a result, it is uncertain when and under what conditions physical risks should be expected. Presumably, text sources seldomly look ahead as far as 2050 or 2100, especially when the intended audience are investors or newspaper readers. This is another important reason why language-based scores may differ from each other and from model-based scores. Identifying the future time horizon and underlying scenario for a specific language-based score would reduce uncertainty in the score estimates and is one next step to increase comparability between scores.

7. Conclusion

Our results show that currently available metrics of firm-level physical climate risk diverge substantially. This divergence raises doubts whether investors have the necessary information to reflect physical climate risks in the prices of corporate stocks and bonds. Absent reliable information, investors are exposed to sudden price adjustments when physical risks materialize, and companies have a limited incentive to invest in climate resiliency.

Financial institutions aiming to integrate physical climate risk into their risk analysis should be aware that the choice of one measure over another is very consequential for the outcome of such an analysis. Therefore, a comprehensive risk management process should avoid placing too much confidence in a single score or ranking and strive to integrate the uncertainties around the assessment of physical risk demonstrated in this study.

Given that more research is needed to drive progress, policymakers can accelerate the development of physical climate risk metrics. Commercial providers of climate risk scores operate in a competitive market, which fosters innovation, but discourages the sharing of underlying data with competitors. Therefore, policymakers should facilitate access to open data by requiring corporations to disclose climate-related risks and relevant performance indicators.⁷ The availability of such data would shift the competitive focus from proprietary data collection to superior data analysis.

Finally, we emphasize that the fact that physical climate risks are hard to estimate does not imply that action should be delayed. There is no doubt that these risks are relevant, as the most recent events in 2021, the heatwave in Canada and the US, and the floods in Europe, have shown. Uncertainty about physical risks increases the urgency to address the underlying drivers of the risk, i.e., to reduce global greenhouse gas emissions.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

⁷ See for example <https://spatialfinanceinitiative.com/geoasset-project/data/>.

Appendix A

A.1. Methodologies

For our analysis, we adopt the widely accepted definitions of physical climate risk provided by the Task Force for Climate-Related Financial Disclosure (TCFD, 2017) and the Sustainability Accounting Standards Board (SASB) (SASB, 2018). The providers of the data used in this study also refer to these definitions. The two definitions state that physical climate risks relate to the exposure of a firm's owned or controlled assets and operations to actual and potential physical impacts of climate change. These impacts can be either acute, i.e., event-driven (e.g., extreme weather), or chronic, i.e., due to long-term shifts in climate patterns (e.g., sea-level rise). Moreover, these risks lead to direct damage to assets, but they also exhibit indirect impacts from supply chain disruption.

Our sample comprises six different measures of physical climate risk. Initially, we identified a total of 19 data sets that offer some form of climate risk scores. Based on a review of methodologies, we narrowed them down to ten data sets which were available for academic research, offered sufficient documentation of their methods, and had the stated goal of assessing physical climate risk. Our final selection of six scores come from both commercial and academic sources and rely on two main approaches:

1. Model-based approaches (labeled as **M1** to **M3**) rely on climate models and combine the model's predictions for various climate hazards with a firm's spatial distribution of assets. Combining these asset-level risks results in an aggregate score for firm-level physical risk (Details of scores are shown in Table A.1).
2. Language-based approaches (labeled **L1** to **L3**) rely on various text sources to collect company-level data concerning physical climate risk topics. **L1** relies on a wide variety of news and social media texts. **L2** relies on earnings call transcripts, and **L3** relies on 10K filings that firms are required to submit to the US Securities and Exchange Commission (SEC). Scores **L2** and **L3** are provided by academics and are available open-source (Sautner et al., 2020b; Leippold, 2020).

A.2. Selection of comparable scores

Our objective is to compare scores that are as consistent with each other as possible. Most providers offer a variety of scores. First, we ensure that all scores reflect the same point in time. As the model-based providers currently only provide data for one point in time, we perform a cross-sectional analysis. It is not possible to obtain all scores at the exact same point in time, but given that physical climate risk is most likely a property that evolves slowly, we do not view this as a major concern. Hence, we take the available figures from each score at the end of 2019, which have been released either in 2018 or 2019. For the language-based scores, which offer time-series data, we take averages over time to create comparable scores. For **L1**, we take the mean for the year 2019 over daily data. For **L2**, we take a four-year average over yearly data covering the years 2016 to 2019. The reason to choose this comparatively long period for **L2** is that this score contains many zeros in any given year. This property is mitigated by averaging over a longer time horizon. For **L3** we choose data based on the 10K filings of 2019.

Second, we ensure that all model-based scores reflect the same modeling choices. This concerns specifically the climate change scenario and the future time horizon. We choose the climate change scenario with RCP8.5 for the main analysis, which is available for **M1** and **M2**. RCP stands for Representative Concentration Pathway, and RCP 8.5 projects the level of radiative forcing by greenhouse gas emissions to stabilize at 8.5 W/m² by 2100. We also provide the corresponding results for RCP4.5. Score **M3** offers temperature scenarios that can be mapped to climate change scenarios, of which the 4-degree warming scenario is comparable to RCP8.5, according to the data provider. For time horizon, we choose 2050 for **M1** and 2040 to 2060 for **M2**. **M3** does not offer scores distinguishing between time horizons. Finally, we use the most aggregated scores from all providers that include all climate hazards rather than scores that focus on specific climate hazards.

The language-based scores do not differentiate between climate change scenarios and future time horizons, but they also come in variations. For **L1** we choose a score called *Insight Score*, the most slow-moving score of three alternatives, which also incorporates information from the past by an exponentially-weighted moving average component. For **L2**, we select the physical risk exposure score. For **L3** we use the physical risk score, computed in the same way as in Kölbel et al. (2020).

Appendix B. Correlations within sectors

We analyze rank correlations within sectors as defined by the Global Industry Classification Standard (GICS). Table B.1 suggests that there are no sectors for which there is a pronounced agreement. On average, we find the highest agreement in *Materials* and lowest in *Real Estate*. The low agreement on *Real Estate* is counterintuitive, since physical risks in real estate are comparatively easier to model and better understood (Baldauf et al., 2020; McNamara and Keeler, 2013).

Among model-based scores, *Information Technology* is the sector where agreement is markedly higher, with significantly positive correlations. However, there are also sectors with significant negative correlations, such as *Financials* and *Consumer Staples*. Out of 33 correlations between model-based scores, 18 are significant, but five of those 18 are negative. Among language-based scores, correlations are mostly insignificant, with the notable exception of *Materials* between L2 and L3 (0.45 with $p < 0.01$). Out of 33 correlations, there are only three significant, none of which is negative. Across language- and model-based approaches, the correlations show no distinctive pattern.

Table A.1

Physical risk scores: Characteristics.

Model-based						
Data Provider	Score	Date	Horizon	Scenario	Hazards	Method description
Trucost, Trucost Climate Change Physical Risk Analytics, M1	Aggregate Score Hazard-specific Scores	2019	Historic Base(2020) 2030 2050	RCP2.6 RCP4.5 RCP8.5	7 Hazards: Water Stress Flood Heatwave Coldwave Hurricane Wildfire Sea Level Rise	5-step procedure: 1. Step: Climate Hazard Mapping 2. Step: Physical Asset Geolocation and Corporate Ownership Mapping 3. Step: Asset and Company Level Physical Risk Scoring 4. Step: Revenue Exposure Based Physical Risk Estimation 5. Step: Composite Score Calculation (Equally or Sensitivity weighted) (Information based on providers' documentation)
Carbon4 Finance, Climate Risk Impact Screening (CRIS), M2	Aggregate Score Hazard-specific Scores	2018	Past Horizon 2040–2070 2070–2100	RCP4.5 RCP6.0 RCP8.5	7 direct Hazards: Acute: changes in the intensity or frequency of heat waves, rainfall extremes and storms, changes in the frequency and duration of drought extremes Chronic: sea-level rise, changes in rainfall patterns, increase in average temperature 9 indirect Hazards: biodiversity migration and loss air quality degradation urban heat island intensification water scarcity wildfires floods (river and groundwater) landslides and mass movements coastal floods coastal erosion	Function of: - Integrated Climate Hazard Rating (Direct Climate Hazard Rating, Risk-aggravating context rating) - Net Vulnerability Rating (Gross-Vulnerability Rating, Adaptability rating) 3-step procedure 1. Step: Collecting geographical and sectoral breakdown of the company's operations 2. Step: Constructing risk ratings for each hazard, each time horizon, and each scenario 3. Step: Calculating Aggregated Risk Ratings - hazard-specific risk scores; arithmetic average - multi-hazard risk scores: geometric average (For details see Carbon 4 (2017))
Southpole, Physical Climate Risk Screening for Investment Portfolios, M3	Aggregate Score	2019	2030–2100	1 °C (RCP2.6) 2 °C (RCP4.5) 3 °C 4 °C (RCP8.5)	Damage functions cover hazards.	Risk is defined via three components: - Exposure - Vulnerability - Hazard 2-step procedure to estimate future physical climate risk for three components: 1. Step: Identification of holding context (geography, sector, etc.) 2. Step: Forward-looking risk assessment (combining climate modeling & vulnerability of company) (Information based on provider's documentation)
Language-based						
Data Provider	Score	Date	Text Source		Method description	
Truvalue Labs, ESG-Scoring (by SASB-Category), L1	Pulse Score Insight Score Momentum Score	2019	Average over daily values		Variety of news and social media sources	
Firm level Climate Change Exposure, Measure of Climate Change Exposure/ Sentiment/ Risk for physical climate risk, L2	Exposure Score Sentiment Score Risk Score	2016–2019	Average over yearly values		Earnings Call Transcripts	
BERT-based climate risk measure L3	Physical Risk Score	2019	One value per year		10-K filings	
					4-step procedure: 1. Step: Data gathering (unstructured data) 2. Step: Natural-Language Processing & AI, Identification of ESG issues and companies 3. Step: Sentiment Analysis, positive/negative ESG-issues 4. Step: Score for certain SASB-category, here: Physical Impacts of Climate Change (Information based on provider's documentation)	
					4-step-procedure: 1. Step: Earnings Call data gathering 2. Step: Pre-construction of bi-grams set relating to physical climate change (ML-Algorithm) 3. Step: Count frequency of bigrams in an earnings call 4. Step: Normalization and eventually transform data from monthly frequency to annual frequency (For more details see Sautner et al. (2020a))	
					3-step procedure: 1. Step: Use BERT to derive probability scores by sentence (item 1A of 10-K-filing) whether it is related to climate, physical or transition risk 2. Step: Create Binary-Score whether probability score by sentence is above a threshold (pre-specified probability threshold: 0.8) 3. Step: Create Score for each file of the company by averaging over binary scores of one document (For more details see Kölbel et al. (2020))	

Table B.1

Kendall's τ by sector: Horizon 2050, RCP8.5.

Score-Pairs	N	Model		Language			Across approaches										Avg
		M1 M2	M1 M3	M2 M3	L1 L2	L1 L3	L2 L3	M1 L1	M1 L2	M1 L3	M2 L1	M2 L2	M2 L3	M3 L1	M3 L2	M3 L3	
Sector																	
Materials	22	0.14	-0.18	0.33	0.16	0.08	0.45	0.27	0.41	0.27	0.09	0.11	0.23	0.15	0.03	0.17	0.18
Industrials	56	0.17	0.11	0.41	0.09	-0.07	-0.05	0.23	-0.09	0.13	0.07	0.06	0.09	0.11	0.14	0.01	0.09
Financials	52	-0.01	0.10	-0.34	0.05	0.22	0.19	0.11	0.34	0.09	0.12	0.17	0.02	-0.10	-0.10	0.14	0.07
Health Care	50	0.00	-0.01	0.35	0.20	-0.06	-0.02	0.08	-0.18	0.10	0.05	-0.08	0.04	0.06	0.20	0.12	0.06
Communication Services	20	-0.12	-0.42	0.34		0.12		-0.17		0.27	0.18		0.19	0.26		-0.20	0.05
Information Technology	56	0.36	0.27	0.45	-0.10	0.09	0.01	-0.03	-0.18	0.00	-0.14	-0.18	0.08	-0.15	-0.02	0.07	0.04
Utilities	24	0.25	-0.01	-0.18	0.13	0.03	-0.04	0.04	0.21	-0.01	0.07	-0.31	0.10	0.03	-0.03	0.21	0.03
Consumer Discretionary	51	-0.06	-0.13	0.18	0.02	-0.13	0.14	-0.10	0.31	0.41	0.16	-0.07	-0.22	0.12	-0.14	-0.19	0.02
Consumer Staples	26	-0.27	-0.36	0.34	0.24	-0.08	-0.08	0.08	-0.24	0.11	-0.03	0.21	-0.05	-0.05	0.21	0.03	0.00
Energy	24	0.26	-0.16	0.18	0.19	0.05	0.08	-0.06	0.32	-0.04	-0.31	0.02	-0.05	-0.42	0.01	-0.16	-0.01
Real Estate	27	-0.22	-0.30	0.28	-0.11	-0.21	0.18	-0.46	0.22	0.21	0.01	-0.19	-0.33	0.28	-0.04	-0.08	-0.05
Avg		0.05	-0.10	0.21	0.09	0.00	0.08	0.00	0.11	0.14	0.02	-0.03	0.01	0.03	0.03	0.01	0.04

This table displays Kendall's τ rank correlations at the GICS sector level for each score pair. Scores M1 to M3 denote model-based scores, L1 to L3 denote language-based scores. ($p < 0.01$ in red, $p < 0.05$ in blue, $p < 0.1$ in green).

References

- Addoum, J.M., Ng, D.T., Ortiz-Bobea, A., 2020. Temperature shocks and establishment sales. *Rev. Financ. Stud.* 33 (3), 1331–1366. <http://dx.doi.org/10.1093/rfs/hhz126>.
- Baldauf, M., Garlappi, L., Yannelis, C., 2020. Does climate change affect real estate prices? Only if you believe in it. *Rev. Financ. Stud.* 33 (3), 1256–1295. <http://dx.doi.org/10.1093/rfs/hhz073>.
- Bansal, R., Kiku, D., Ochoa, M., 2016. Price of long-run temperature shifts in capital markets. Working Paper w22529. National Bureau of Economic Research, Cambridge, MA. <http://dx.doi.org/10.3386/w22529>.
- Behnam, R., Litterman, B., Gillers, D., Martinez-Diaz, L., Keenan, J.M., Moch, S., 2020. Managing climate risk in the U.S. financial system. Technical Report.
- Bolton, P., Kacperczyk, M.T., 2020. Global pricing of carbon-transition risk. Working Paper. <http://dx.doi.org/10.2139/ssrn.3550233>.
- Brock, W.A., Hansen, L.P., 2018. Wrestling with uncertainty in climate economic models. University of Chicago, Becker Friedman Institute for Economics Working Paper 2019-71. <http://dx.doi.org/10.2139/ssrn.3008833>.
- Carbon 4, 2017. CRIS Climate risk impact screening, a unique method to assess the impacts of physical risks from climate change on financial assets, the methodological guidebook. Technical Report. http://crisforfinance.com/wp-content/uploads/2017/11/CRIS-Guidebook_Publicversion_Nov2017.pdf.
- Fiedler, T., Pitman, A.J., Mackenzie, K., Wood, N., Jakob, C., Perkins-Kirkpatrick, S.E., 2021. Business risk and the emergence of climate analytics. *Nature Clim. Change* 11 (2), 87–94.
- Herring, S.C., Hoerling, M.P., Kossin, J.P., Peterson, T.C., Stott, P.A., 2015. Explaining extreme events of 2014 from a climate perspective. *Bull. Am. Meteorol. Soc.* 96 (12), 1–172. <http://dx.doi.org/10.1175/BAMS-ExplainingExtremeEvents2014.1>.
- Horwitz, S., 2009. Wal-mart to the rescue: Private enterprise's response to Hurricane Katrina. *The Indep. Rev.* 13 (4), 511–528.
- IPCC, 2014. Climate change 2014: Synthesis Report. Contribution of working groups I, II and III to the fifth assessment report of the intergovernmental panel on climate change. Technical Report ISBN: 9789291691432, pp. 151.
- Kölbel, J.F., Leippold, M., Rillaerts, J., Wang, Q., 2020. Ask BERT: How regulatory disclosure of transition and physical climate risks affects the CDS term structure. Working Paper. <http://dx.doi.org/10.2139/ssrn.3616324>.
- Leippold, M., 2020. Transition and physical climate risk scores based on 10-k filings. data retrieved from <https://osf.io/pk2u9/>.
- McNamara, D.E., Keeler, A., 2013. A coupled physical and economic model of the response of coastal real estate to climate risk. *Nature Clim. Change* 3 (6), 559–562. <http://dx.doi.org/10.1038/nclimate1826>.
- SASB, 2018. SASB Materiality Map. Sustainability Accounting Standards Board.
- Sautner, Z., van Lent, L., Vilkov, G., Zhang, R., 2020a. Firm-level climate change exposure. Working Paper. <http://dx.doi.org/10.2139/ssrn.3642508>.
- Sautner, Z., van Lent, L., Vilkov, G., Zhang, R., 2020b. Firm-level climate change exposure. data retrieved from <https://osf.io/fd6jq/>.
- Stott, P., 2016. How climate change affects extreme weather events. *Science* 352 (6293), 1517–1518. <http://dx.doi.org/10.1126/science.aaf7271>.
- Sun, Y., Zhang, X., Zwiers, F.W., Song, L., Wan, H., Hu, T., Yin, H., Ren, G., 2014. Rapid increase in the risk of extreme summer heat in eastern China. *Nature Clim. Change* 4 (12), 1082–1085. <http://dx.doi.org/10.1038/nclimate2410>.
- Swiss Re Institute, 2020. Natural catastrophes in times of economic accumulation and climate change. *Sigma* 2, 1–32, <https://www.swissre.com/institute/research/sigma-research/sigma-2020-02.html>.
- TCFD, 2017. Recommendations of the task force on climate-related financial disclosures. Technical Report. p. 74.
- TCFD, 2019. Task force on climate-related financial disclosures: Status report. Technical Report. p. 144.
- Varini, F.S., Boyd-Graber, J., Ciaramita, M., Leippold, M., 2020. CLIMATE: A dataset for climate change topic detection. Tackling Climate Change with Machine Learning (Climate Change AI) Workshop at NeurIPS, December 2020.