Cureus Journal of Computer Science
https://doi.org/10.7759/s44389-026-00359-6

Original Article

Does Earnings-Call Climate Exposure Improve Machine-Learning Forecasts of Stock Downside Risk? Out-of-Sample Evidence from U.S. Firms

Nandini Bhati [1], Abhinav Kataria [1], Pankaj Dhaundiyal [1]

1. *School of Business and Management, Christ University, Bengaluru, IND*

Received: September 27, 2026 | Review began: October 01, 2026 | Review ended: October 08, 2026 | Published: November 06, 2026

© **Copyright** 2026
This is an open access article distributed under the terms of the Creative Commons Attribution License CC-BY 4.0., which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

Climate-related risks are now widely regarded as financial risks, and machine learning is often proposed as a means of turning climate information into sharper risk assessments. Whether firm-level climate exposure helps forecast a stock's future risk, rather than merely correlating with it in-sample, has received much less attention. We examine this question for 1,792 U.S.-listed non-financial firms (15,521 firm-years, with portfolio formation dates from 2012 to 2024), combining published earnings-call measures of firm-level climate change exposure with freely available accounting and market data. Our forecast target is each firm's downside semi-deviation over the following year. We estimate ordinary least squares and random forest models without climate exposure, with it, and with placebo climate variables randomly reassigned among firms in the same year and industry. We evaluate the forecasts in out-of-sample tests over 2017-2024 (10,836 firm-years) and use them to form sector-neutral portfolios that exclude the stocks with the highest predicted risk. Climate exposure does not improve forecast accuracy: adding it moves the random forest's out-of-sample R^2 from 10.25% to 10.30% (Diebold-Mariano $p = 0.665$), and the model with real climate data performs no better than 100 placebo versions (permutation $p = 0.168$). Random forests are more accurate than linear models ($p = 0.022$), but this advantage is only marginally significant after a small-sample correction ($p = 0.069$), and most of their gain over a naive past-risk forecast arises in the year after the COVID-19 shock; excluding that year, the pooled out-of-sample R^2 is −3.3% for the random forest and −8.4% for the linear model. Portfolios formed with and without climate information have almost identical risk. For U.S.-listed firms and a one-year horizon, earnings-call climate exposure therefore adds little to forecasts of downside risk once past volatility and basic firm characteristics are known, and the contributions of climate data and of model flexibility should be assessed separately.

Categories: AI applications, Predictive Analytics, Machine Learning (ML)

Keywords: climate risk, climate change exposure, machine learning, random forest, downside risk, volatility forecasting, portfolio management, sustainable finance

Introduction

The Network for Greening the Financial System, which brings together central banks and financial supervisors, regards climate-related risks as a source of financial risk that falls within its members' mandates, and it has asked for better data to measure them [1]. The climate finance literature separates physical risks, which stem from the direct effects of a changing climate, from transition risks, which stem from the shift to a low-carbon economy, and studies how each shows up in asset prices and in the construction of hedging portfolios [2]. Moving from these economy-wide risks to their

consequences for individual firms is considerably harder, as work on climate stress testing makes clear [3]. A portfolio manager faces a narrower and more practical question: if we know how exposed a firm is to climate change, can we better anticipate how risky its stock will be?

Most of the evidence on that question comes from in-sample analysis. Option prices show that insurance against extreme price falls costs more for carbon-intensive firms, and more so when public attention to climate change is high [4]. Sautner et al. [5] derive firms' climate change exposure from the text of their earnings conference calls and show that this information is priced in both option and equity markets. Using the same measures for firms in 54 countries, Naseer et al. [6] find that firms with higher exposure have more volatile stocks and that the relationship is driven by regulatory and opportunity exposure rather than by physical exposure.

An association between climate exposure and risk does not by itself establish that exposure helps forecast risk. Because volatility is highly persistent, a variable can carry a significant coefficient in a full-sample regression and still contribute nothing to a forecast once recent volatility is included. Its apparent contribution may also disappear when a model uses only past data for estimation and faces evaluation on subsequent years. Investors must form expectations before outcomes materialize, and the out-of-sample test is therefore the relevant one for them.

Machine learning, now applied to a wide range of climate-related problems [7], raises a second and distinct question. In a large comparison of methods for forecasting stock returns, tree-based models and neural networks achieved the highest accuracy, and the authors traced these gains to nonlinear interactions among predictors [8]. The evidence for volatility forecasting is more mixed: machine-learning models outperform standard econometric benchmarks, but the random forest's margin over a regularized linear model is small and usually not statistically significant [9]. Studies that combine machine learning with climate variables have mostly forecast a single index or a market-level series [10,11], or have constructed small portfolios whose objective is not risk prediction [12]. When a machine-learning model with climate inputs performs well, it is therefore difficult to determine whether the improvement stems from the climate data, from the flexibility of the model, or from both.

This study separates these two sources. The forecast target, for each firm and year, is next-year downside semi-deviation, the component of return volatility attributable to price declines. Both ordinary least squares and a random forest are estimated with and without the exposure measures of Sautner et al. [5]. Placebo versions of the climate models, in which exposure values are randomly reassigned among firms in the same year and industry, reveal whether any improvement reflects firm-specific climate information or merely the addition of predictors. The sample, constructed entirely from free sources, comprises 1,792 U.S.-listed non-financial firms and 15,521 firm-years. Forecasts are generated in an expanding window and evaluated over 2017-2024, and we use them to form sector-neutral portfolios that exclude the stocks with the highest predicted risk.

We find that climate exposure contributes no meaningful predictive power. Relative to a naive forecast that uses last year's downside risk, the random forest reaches an out-of-sample R^2 of 10.25% without climate data and 10.30% with it. The difference is not statistically significant, and the model using the real climate data does no better than its placebo counterparts. The random forest is more accurate than the linear model, although with eight test years this advantage is only marginally significant after a small-sample correction, and most of its gain over the naive forecast comes from the year after the COVID-19 shock. Portfolios built from forecasts with and without climate information differ very little in risk.

The paper makes three contributions. First, it provides an out-of-sample, firm-level test of whether climate exposure measured from earnings calls improves forecasts of downside risk, complementing the in-sample evidence [6] and the market-level forecasting evidence [11,13]. Second, its design distinguishes the value of climate information from the value of a more flexible model, a distinction absent from the studies reviewed here that combine machine learning with climate data. Third, it documents what free data can achieve and where they fall short, particularly survivorship bias at the identifier-matching stage and share counts that require correction before market values can be used.

How to cite this article:
Bhati N, Kataria A, Dhaundiyal P (November 06, 2026) Does Earnings-Call Climate Exposure Improve Machine-Learning Forecasts of Stock Downside Risk? Out-of-Sample Evidence from U.S. Firms. Cureus J Comput Sci 3 : es44389-026-00359-6.
DOI https://doi.org/10.7759/s44389-026-00359-6

Background and literature

Four bodies of research are relevant here: studies of whether climate risk is priced, of how firm-level climate exposure can be measured, of the relationship between climate exposure and return volatility and of machine learning in financial forecasting. We take them one at a time before turning to the points at which they overlap.

Climate Risk and Asset Prices

Climate change can affect financial markets in two broad ways. Physical risks, such as drought, flooding and heat, damage firms' operations and assets, while transition risks arise from the changes in policy, technology and demand that accompany the move to a low-carbon economy [1,2]. A considerable body of evidence now indicates that equity markets price at least part of these risks. Bolton and Kacperczyk [14] show that U.S. firms with higher carbon emissions, and with faster growth in emissions, earn higher stock returns after standard controls, which they interpret as compensation for carbon risk. Pástor et al. [15] present an equilibrium model in which investors' preference for sustainable firms shapes the expected returns of green and brown assets.

There is less agreement about which climate risks are priced. Faccini et al. [16] build four market-wide climate risk factors from a textual analysis of Reuters news on climate change and find that only the factor tied to U.S. climate policy is priced, especially after 2012. For physical risk, Hong et al. [17] show that long-run drought trends across 31 countries predict weaker profit growth and lower stock returns for food companies, which is consistent with markets underreacting to climate information. Read together, these studies suggest that information about the transition, and about policy risk in particular, shows up in prices more clearly than information about physical risk.

Measuring Firm-Level Climate Exposure

Because emissions data are incomplete and data vendors frequently estimate them, a growing literature measures climate exposure from what firms and analysts say. Sautner et al. [5] apply a machine-learning keyword-discovery algorithm to earnings-call transcripts and construct quarterly measures of overall, opportunity, regulatory and physical exposure for more than 10,000 firms. Kölbel et al. [18] fine-tune the BERT (Bidirectional Encoder Representations from Transformers) language model to distinguish transition-risk from physical-risk disclosures in U.S. annual reports and find that, after the Paris Agreement, the two affect credit default swap spreads in opposite directions. Bingler et al. [19] apply ClimateBERT, a BERT model adapted to climate text, to the climate-risk disclosures of firms that support the Task Force on Climate-related Financial Disclosures, and conclude that much of this reporting is cheap talk in which firms selectively disclose non-material information. Researchers also use machine learning to fill gaps in emissions data directly: Nguyen et al. [20] combine several learners to estimate emissions for firms that do not disclose them. Much of the climate information available to investors is thus itself the output of text-mining or machine-learning methods, and its measurement quality becomes part of any analysis that relies on it.

Climate Exposure, Volatility and Downside Risk

A related line of work studies the variability of returns rather than their average. At the firm level, Ilhan et al. [4] show that option-implied protection against extreme price falls costs more for carbon-intensive firms. Naseer et al. [6] report a positive association between the exposure measures of Sautner et al. [5] and annual stock volatility in 54 countries, using fixed-effects, generalized method of moments and instrumental-variable estimators; about 84% of their observations come from North America. At the market level the evidence extends to forecasting. Zhou and Ma [13] find that a composite climate-risk index predicts stock-market volatility in 32 countries, including in out-of-sample tests, and Bonato et al. [11] use a machine-learning bagging method to show that climate risks help forecast realized volatility in U.S. states one to two months ahead. Salisu et al. [21] ask whether uncertainty in temperature growth helps forecast tail risk in the U.S. stock market, using data that go back to 1895. In short, the link between climate exposure and risk has been documented in-sample for firms and in out-of-sample tests for markets. Whether firm-level exposure improves out-of-sample forecasts of firm-level risk is a separate question that remains open.

Machine Learning for Financial Forecasting

How to cite this article:
Bhati N, Kataria A, Dhaundiyal P (November 06, 2026) Does Earnings-Call Climate Exposure Improve Machine-Learning Forecasts of Stock Downside Risk? Out-of-Sample Evidence from U.S. Firms. Cureus J Comput Sci 3 : es44389-026-00359-6.
DOI https://doi.org/10.7759/s44389-026-00359-6

Gu et al. [8] compare linear models, penalized regressions, dimension-reduction methods, boosted trees, random forests and neural networks for forecasting stock returns; evaluate them in strictly out-of-sample tests; and find that trees and neural networks perform best. The random forest is among the most widely used of these methods; it averages many decision trees, each grown on a bootstrap sample of the data, with a random subset of predictors considered at every split [22]. For volatility, Christensen et al. [9] find that machine-learning models outperform standard econometric benchmarks for large U.S. stocks, although the random forest's advantage over a regularized linear model is small and seldom statistically significant. Validation procedures matter in such comparisons. Assael et al. [23], for example, apply gradient boosting to European equities with environmental, social and governance (ESG) features but validate with random cross-validation instead of a time-ordered split, which can allow information from later periods to enter training. Comparative studies of evaluation schemes indicate that, when data are non-stationary, validation that preserves the temporal order of observations yields the most accurate estimates of forecast performance [24]. SHAP (SHapley Additive exPlanations) values [25] quantify the contribution of each input to a tree model's predictions, and for tree ensembles they can be computed exactly and efficiently [26].

Where the Strands Meet

The studies closest to ours bring together machine learning, climate variables and investment outcomes, but each aims at something different. Awijen et al. [10] show that transition and physical climate-risk variables improve machine-learning predictions of the S&P 500 ESG index and interpret their models with SHAP; what they forecast is a single index price, not the risk of individual firms. Bouyaddou and Jebabli [12] build a reinforcement-learning portfolio of 30 Dow Jones stocks that trades off returns against the portfolio's carbon footprint, so carbon enters as an objective rather than as a predictor of financial risk. Engle et al. [27] form portfolios that hedge innovations in a newspaper-based climate news index, measuring firms' exposure with third-party ESG scores; their goal is to hedge climate news, not to forecast firm-level downside risk. None of these studies asks whether firm-level climate exposure improves out-of-sample risk forecasts once conventional predictors are in the model, or whether an improvement should be credited to the data or to the model.

Research gap

Three gaps emerge from this review. The association between firm-level climate exposure and stock volatility has been shown in-sample [6], but the studies reviewed do not test whether that exposure improves out-of-sample forecasts of firm-level risk. Where machine learning and climate data have been used together, the unit of analysis is an index or a market [10,11], and the contribution of the climate data is not isolated from that of a more flexible model. Finally, none of the studies we reviewed turns firm-level, climate-informed risk forecasts into benchmark-relative portfolios and evaluates them in out-of-sample tests. Because our literature search was targeted rather than systematic, these gaps describe the studies reviewed here and not necessarily the whole field. We address them using free data only, which limits the sample to U.S.-listed firms.

Research questions

These gaps lead to four research questions:

RQ1. Does firm-level climate change exposure improve out-of-sample forecasts of next-year downside risk beyond conventional risk predictors?

RQ2. Does a random forest extract more predictive information than a linear model from the same inputs?

RQ3. Do portfolios that use climate-informed forecasts to exclude high-risk stocks differ in realized risk from portfolios that use forecasts without climate information?

RQ4. Which dimension of exposure (regulatory, physical or opportunity) drives any predictive contribution?

How to cite this article:
Bhati N, Kataria A, Dhaundiyal P (November 06, 2026) Does Earnings-Call Climate Exposure Improve Machine-Learning Forecasts of Stock Downside Risk? Out-of-Sample Evidence from U.S. Firms. Cureus J Comput Sci 3 : es44389-026-00359-6.
DOI https://doi.org/10.7759/s44389-026-00359-6

Materials And Methods

Conceptual framework and hypotheses

Figure *1* outlines three routes by which climate exposure might increase a firm's downside risk. Regulatory exposure makes future compliance costs and the value of carbon-intensive assets uncertain, a risk that option markets appear to price [4]. Opportunity exposure points to investment in new technologies whose payoffs are uncertain [6]. Physical exposure affects operations and supply chains, although the evidence suggests that markets price it only weakly [16,17]. If these routes are at work, climate exposure should carry information about future downside risk. The open question is whether that information goes beyond what past volatility already reveals: since volatility is persistent, climate-related risk that has already affected a stock may already be reflected in its recent volatility.

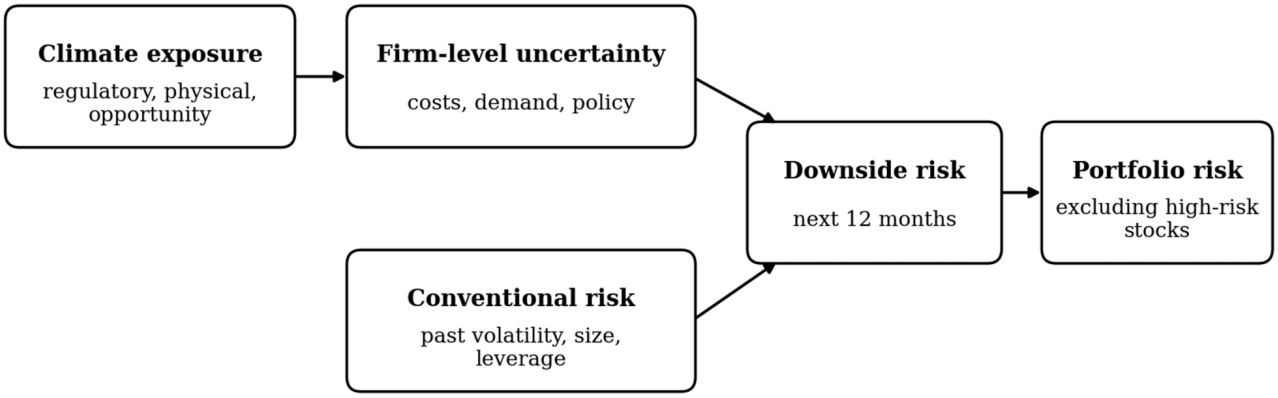


FIGURE 1: Conceptual framework

The empirical tests concern the path from climate exposure to next-year downside risk after conditioning on conventional risk measures and the effect of risk forecasts on portfolio risk.

H1 (climate information; RQ1). Out-of-sample forecast error for next-year downside risk falls when firm-level climate exposure is added to the conventional predictors. This hypothesis rests on the in-sample association reported by Naseer et al. [6], on the pricing of carbon tail risk [4] and on predictability at the market level [13]. The competing expectation is that nothing is gained, because past volatility already reflects climate-related risk.

H1b (dimension; RQ4). Any predictive contribution comes mainly from regulatory rather than physical exposure, in line with evidence that markets price policy-related climate risk more reliably than physical risk [6,16].

H2 (model flexibility; RQ2). Given the same inputs, a random forest and a linear model differ in forecast accuracy. We do not specify the direction in advance, since nonlinear models improve return prediction [8] but have produced only small gains in volatility forecasting [9].

H3 (economic value; RQ3). When climate-informed forecasts are used to exclude the stocks with the highest predicted risk, the resulting portfolios have lower realized downside risk than portfolios built from forecasts without climate information. We expect this only if H1 holds.

Data sources

Every input is free and publicly available (Table *1*). We did not have access to the commercial databases common in empirical finance, such as Center for Research in Security Prices, Compustat, Bloomberg and Morgan Stanley Capital International, and none of them is used. Several features of the sample described below follow from this constraint.

How to cite this article:
Bhati N, Kataria A, Dhaundiyal P (November 06, 2026) Does Earnings-Call Climate Exposure Improve Machine-Learning Forecasts of Stock Downside Risk? Out-of-Sample Evidence from U.S. Firms. Cureus J Comput Sci 3 : es44389-026-00359-6.
DOI https://doi.org/10.7759/s44389-026-00359-6

Source	Data used	Identifiers
Sautner et al. [28], OSF repository; file firmquarter_score_2024Q4_Version_2025_Jul_03.csv	Quarterly equal-weighted climate change exposure (overall, regulatory, physical, opportunity), 2001–2024; 427,467 firm-quarters for 15,928 firms	ISIN, GVKEY, CUSIP, headquarters country
U.S. Securities and Exchange Commission [29], EDGAR application programming interfaces	Form 10-K accounting items with filing dates; shares outstanding from 10-K and 10-Q cover pages; SIC industry code	CIK, ticker
OpenFIGI [30] mapping service	Conversion of ISINs to U.S. common-stock tickers	FIGI, ticker
Yahoo Finance, via the open-source yfinance package (version 0.2.66); Aroussi [31]	Daily closing prices, dividend-adjusted prices and stock splits, January 2011–July 2025; SPY exchange-traded fund as market proxy	Ticker
French [32], Data Library	Monthly risk-free rate	—

TABLE 1: Data sources

GVKEY refers to the Compustat company key provided in the climate file.
CIK, SEC Central Index Key; CUSIP, Committee on Uniform Securities Identification Procedures; EDGAR, Electronic Data Gathering, Analysis, and Retrieval; FIGI, Financial Instrument Global Identifier; ISIN, International Securities Identification Number; OSF, Open Science Framework; SEC, U.S. Securities and Exchange Commission; SIC, Standard Industrial Classification; SPY, SPDR S&P 500 ETF Trust

Sample construction

The starting universe consists of every firm in the climate file with a U.S. International Securities Identification Number or U.S. headquarters. Firms must file annual reports with the U.S. Securities and Exchange Commission (SEC) on Form 10-K, and financial firms (Standard Industrial Classification [SIC] 6000-6799) and public-administration entities are excluded. Portfolios are formed on 30 June of each year from 2012 to 2024. A firm-year enters the sample if the firm has at least 200 daily returns over the preceding year, a price within 10 days of the formation date, at least two earnings calls in the climate measurement window and a valid market value. Table *2* traces the sample through each step.

Step	Firms	Firm-years
Climate file, all firms	15,928	—
U.S. ISIN or U.S. headquarters	8,227	—
With climate exposure at one or more formation dates	6,279	41,415

How to cite this article:
Bhati N, Kataria A, Dhaundiyal P (November 06, 2026) Does Earnings-Call Climate Exposure Improve Machine-Learning Forecasts of Stock Downside Risk? Out-of-Sample Evidence from U.S. Firms. Cureus J Comput Sci 3 : es44389-026-00359-6.
DOI https://doi.org/10.7759/s44389-026-00359-6

Matched to a current U.S. common-stock ticker	2,490	—
Matched to an SEC CIK	2,318	—
With SEC accounting data	2,312	—
With daily price history	2,316 of 2,318	—
Investable at formation (10-K filer, return history, price)	2,301	23,498
Merged with climate exposure	2,298	20,296
Excluding financial firms, public administration and missing industry codes	1,821	16,111
With a valid market value (final sample)	1,792	15,521
Out-of-sample test years, 2017–2024	—	10,836

TABLE 2: Sample construction

Counts of firms after identifier matching refer to the matched set; the price-history count is out of the 2,318 firms matched to a CIK. CIK, Central Index Key; ISIN, International Securities Identification Number; SEC, U.S. Securities and Exchange Commission

Most attrition occurs at the ticker-matching stage: only 2,490 of the 6,279 U.S. firms with climate data could be matched to a current ticker. For 70% of the unmatched firms, the climate file contains no earnings call after 2022, which suggests that most of them were delisted or acquired. A further 121 firms were listed in the United States only through American depositary receipts (ADRs) and are excluded for the reasons explained below. Once matched, almost no firms drop out at the price stage (2 of 2,318).

Identifier matching and market values

The datasets are linked through exact identifiers only; company names play no role in the matching. The OpenFIGI service converts International Securities Identification Numbers from the climate file to U.S. common-stock tickers, restricted to common-stock and Real Estate Investment Trust listings. Tickers are standardized (e.g., “BRK/B” becomes “BRK-B”) and matched to the SEC's list of current tickers, which provides each firm's SEC Central Index Key and, through it, the firm's accounting data. An accounting value enters the analysis only if it was filed on or before the formation date and refers to a fiscal period that ended no more than about 18 months earlier. Where total liabilities are not reported, they are computed as total assets minus shareholders' equity (5,348 firm-years).

Market capitalization equals the unadjusted closing price on the formation date multiplied by the number of shares outstanding. Share counts come from the cover pages of 10-K and 10-Q filings and are adjusted for any stock split between the filing date and the formation date, so that price and share count refer to the same share basis. Where no cover-page count exists, the diluted weighted-average share count from the annual report serves as a substitute (1,855 firm-years). ADR listings are excluded because their prices refer to depositary shares, whereas SEC share counts refer to ordinary shares. Because some filings report share counts on the wrong scale, market value is set to missing in 567 firm-years (258 firms) in which the split-adjusted share count lies more than ten times away from the firm's own median, the

How to cite this article:
Bhati N, Kataria A, Dhaundiyal P (November 06, 2026) Does Earnings-Call Climate Exposure Improve Machine-Learning Forecasts of Stock Downside Risk? Out-of-Sample Evidence from U.S. Firms. Cureus J Comput Sci 3 : es44389-026-00359-6.
DOI https://doi.org/10.7759/s44389-026-00359-6

cover-page and weighted-average counts disagree by more than a factor of 10, or the implied market value exceeds $4 trillion. These screens affect none of the largest firms, and, as the Results and Discussion section shows, the resulting value-weighted benchmark closely tracks the U.S. market.

Timing of information

Every predictor uses only information that was public on the formation date, and accounting data enter only after their filing date. Climate exposure is the average over the four quarterly earnings calls ending in the third quarter of the previous calendar year; if the quarters are calendar quarters, the measure is therefore at least nine months old at formation. This conservative window was chosen because the data documentation does not state whether the quarter field refers to calendar or fiscal quarters, and under either reading every call in the window took place before 30 June. Past-return measures end on the formation date, and the forecast target begins on the next trading day. One limitation remains: Sautner et al. [5] learned their keyword list from transcripts spanning the entire sample period, so the measure is not strictly real-time, even though each call was public when it took place.

Variables

Table *3* defines the variables. The forecast target, specified in detail below, is next-year downside semi-deviation. The conventional predictors capture past risk (volatility, downside semi-deviation and market beta), past performance (momentum), firm characteristics (size, book-to-market and leverage) and industry. The climate predictors are the four exposure measures of Sautner et al. [5]. To limit the influence of extreme values, all predictors are winsorized at the 1st and 99th percentiles within each year. The analysis uses no emissions data, because no free source providing point-in-time emissions for firms across all industries could be identified.

Variable	Definition	Source
Downside semi-deviation, next 12 months (target)	Annualized square root of the mean squared negative daily return, July–June after formation; used in logs	Yahoo Finance
Volatility, next 12 months (alternative target)	Annualized standard deviation of daily returns, July–June after formation	Yahoo Finance
Log past volatility	Log of the annualized standard deviation of daily returns in the year before formation	Yahoo Finance
Log past downside semi-deviation	As the target, measured over the year before formation	Yahoo Finance
Beta	Sensitivity of daily returns to SPY returns in the year before formation	Yahoo Finance
Momentum	Return from 12 months to 1 month before formation	Yahoo Finance
Log market capitalization	Log of price × shares outstanding at formation	Yahoo Finance; SEC
Book-to-market	Shareholders' equity divided by market capitalization	SEC; Yahoo Finance

How to cite this article:
Bhati N, Kataria A, Dhaundiyal P (November 06, 2026) Does Earnings-Call Climate Exposure Improve Machine-Learning Forecasts of Stock Downside Risk? Out-of-Sample Evidence from U.S. Firms. Cureus J Comput Sci 3 : es44389-026-00359-6.
DOI https://doi.org/10.7759/s44389-026-00359-6

Leverage	Total liabilities divided by total assets	SEC
Industry	SIC division	SEC
Overall, regulatory, physical and opportunity exposure	Mean of the quarterly equal-weighted exposure measures over the climate window	Sautner et al. [28]

TABLE 3: Variable definitions

SEC, U.S. Securities and Exchange Commission; SIC, Standard Industrial Classification; SPY, SPDR S&P 500 ETF Trust

Forecast target

Let *t* denote a formation date (30 June of each year). For firm *i*, downside semi-deviation over the following 365 days is computed from the firm's *N* daily returns *r* in that window, with at least 60 returns required:

$$DSD_{i,t} = \sqrt{252} \cdot \left(\frac{1}{N}\sum_{d=1}^{N} \min(r_{i,d}, 0)^2\right)^{1/2}, \qquad y_{i,t} = \ln DSD_{i,t} \tag{1}$$

Ordinary volatility treats rises and falls alike, whereas this measure counts only falls, which comes closer to investors' concern with losses. Taking the logarithm reduces the skewness of the measure. Realized volatility over the same period serves as an alternative target.

Forecasting models

Every model draws on one of three predictor sets: the conventional predictors X on their own, X together with the climate exposure measures C or X together with placebo climate variables $\tilde{C}$ (Table *4*). The naive benchmark simply carries last year's downside semi-deviation forward. Because past volatility and past downside semi-deviation enter the models in logarithms, the linear model can reproduce the naive forecast as a special case.

Model	Specification
Naive	Forecast equals the log of last year's downside semi-deviation
M1 / M2 / M2p	Ordinary least squares on X / X + C / X + $\tilde{C}$
M3 / M4 / M4p	Random forest on X / X + C / X + $\tilde{C}$

TABLE 4: Forecasting models

X denotes conventional predictors, C denotes climate exposure measures and $\tilde{C}$ denotes placebo climate variables.

A random forest averages the predictions of many decision trees. Each tree grows on a bootstrap sample of the observations and considers only a random subset of predictors at each split [22]. Because every tree partitions the data at threshold values of the predictors, the forest can capture nonlinear effects and interactions without prior specification; climate exposure might, for instance, matter only for firms with low past volatility. Four considerations favor the random

How to cite this article:
Bhati N, Kataria A, Dhaundiyal P (November 06, 2026) Does Earnings-Call Climate Exposure Improve Machine-Learning Forecasts of Stock Downside Risk? Out-of-Sample Evidence from U.S. Firms. Cureus J Comput Sci 3 : es44389-026-00359-6.
DOI https://doi.org/10.7759/s44389-026-00359-6

forest over more complex methods. It performed best among tree-based methods in a comparable study of volatility forecasting [9]. It usually performs well with little tuning [33], which limits overfitting in an annual panel. SHAP values can decompose its predictions [25]. With at most 13 annual observations per firm, the panel is too short for recurrent neural networks. No new language model is trained, because the climate measures are themselves the output of a validated text-mining method [5].

The placebo variables address a specific concern. Adding columns to a random forest alters how it samples predictors, so its accuracy can change even when the new columns carry no information. For each placebo, we randomly reassign the four climate variables among firms in the same year and SIC division; this preserves their distribution and industry-level information but breaks the link to individual firms. We draw 100 independent placebos for each model. Because the random forest is far more costly to estimate, each random-forest placebo uses the hyperparameters selected for the real climate model in the same test year instead of being tuned separately. Reassignment also removes any correlation between climate exposure and other firm characteristics, so the placebo test asks whether the climate variables are more useful than random numbers. The direct comparisons of M2 with M1 and of M4 with M3 answer the separate question of whether they add information beyond the other predictors.

Training and validation

Forecasts are generated with an expanding window. For each test year T from 2017 to 2024, the models are trained on all earlier firm-years whose outcomes were fully observed by 30 June of year T and then applied to every firm that is investable in year T. Observations are never shuffled across time, so no information from the test year enters training. The random forest's settings are selected with the final training year alone as the validation set: the minimum number of observations in a terminal leaf (20, 50 or 100) and the share of predictors considered at each split (0.33 or 0.66). Tuning uses 200 trees and the final model 300. Missing predictor values are replaced with training-sample medians, and all predictors are standardized with training-sample means and standard deviations. Industry enters as SIC-division indicator variables. Estimation is performed in Python 3.13 with scikit-learn 1.6.1 [34], statsmodels 0.15.0 [35] and shap 0.52.0 [25], using a fixed random seed (42); the list of package versions accompanies the code [36].

Forecast evaluation

We measure forecast accuracy by the root mean squared error (RMSE), the mean absolute error (MAE) and the out-of-sample R^2 relative to the naive forecast:

$$R^2_{OOS} = 1 - \frac{\sum_{i,t}(y_{i,t} - \hat{y}_{i,t})^2}{\sum_{i,t}(y_{i,t} - \hat{y}_{naive,i,t})^2} \tag{2}$$

A positive value indicates that the model outperforms the naive forecast. We compare pairs of models through the difference in their mean squared errors (MSE) in two ways: a 95% interval from a bootstrap that resamples firms (1,000 draws), and a Diebold-Mariano test [37] applied to the eight annual averages of the loss difference, with a Newey-West variance using one lag. The firm bootstrap does not capture the tendency of forecast errors for different firms to move together within a year, so where the two approaches disagree, we treat the annual test as the more conservative guide. Because the annual test rests on only eight observations, we also report it with a small-sample correction (Table *9*).

For the climate models, the permutation p-value is the share of placebos that perform at least as well as the real climate data:

$$p = \frac{1 + \#\{\text{placebos with MSE} \le \text{MSE}_{real}\}}{1 + n_{placebo}} \tag{3}$$

Econometric benchmark

To link the forecasting results with earlier in-sample evidence, we estimate a panel regression of the target on the climate measures C and the conventional predictors X, with year effects α, industry effects γ and standard errors clustered by firm:

How to cite this article:
Bhati N, Kataria A, Dhaundiyal P (November 06, 2026) Does Earnings-Call Climate Exposure Improve Machine-Learning Forecasts of Stock Downside Risk? Out-of-Sample Evidence from U.S. Firms. Cureus J Comput Sci 3 : es44389-026-00359-6.
DOI https://doi.org/10.7759/s44389-026-00359-6

$$y_{i,t} = \alpha_t + \gamma_{s(i)} + \beta' C_{i,t} + \delta' X_{i,t} + \varepsilon_{i,t} \quad (4)$$

As a further check, we run annual cross-sectional regressions and average their coefficients in the manner of Fama and MacBeth [38], with Newey-West standard errors using one lag. These regressions describe associations; they are not designed to identify causal effects.

Portfolio construction

On each formation date in the test period, stocks ranked in the top 20% of the relevant risk score within their SIC division are excluded, and divisions with fewer than five stocks remain unscreened. The remaining stocks are value-weighted by market capitalization and held for 12 months without interim rebalancing (Table *5*). Because the exclusion operates within industries, the portfolios remain approximately sector-neutral, so performance differences between them cannot be attributed to industry tilts. The investment universe includes every firm investable on the formation date, regardless of whether it survives the holding period. Free data sources do not provide delisting returns, so a stock that stops trading is treated as cash for the remainder of the year. Transaction costs of 10 basis points per side apply to turnover at each annual rebalancing.

Portfolio	Score used to exclude stocks
P0	None: value-weighted benchmark of the whole universe
P1	Regulatory climate exposure (a simple screen without a model)
P2 / P3	Random-forest forecast without / with climate data (M3 / M4)
P2L / P3L	Linear-model forecast without / with climate data (M1 / M2)

TABLE 5: Portfolios

Models M1–M4 are defined in Table 4. Each portfolio other than P0 excludes the stocks in the top 20% of its score within each SIC division, and the remaining stocks are weighted by market value.
SIC, Standard Industrial Classification

Performance is evaluated over the 96 monthly returns from July 2017 to June 2025. The reported measures are annualized return and volatility; Sharpe and Sortino ratios computed from returns in excess of the risk-free rate [32]; maximum drawdown; the historical 95% value at risk and expected shortfall of monthly returns; tracking error and the information ratio relative to P0; and turnover. Differences in Sharpe ratios are tested with the method of Ledoit and Wolf [39], which accommodates non-normal and autocorrelated returns. To validate the reconstructed market values, the benchmark P0 is compared with the SPY (SPDR S&P 500 ETF Trust) market fund over the same months.

Interpretation and robustness

Mean absolute SHAP values measure how much each predictor contributes to the predictions of the climate-informed random forest, computing them separately for each of the eight test years and reporting their average and range. All robustness checks were specified before any results were examined. They use realized volatility as the target; replace the random forest with gradient boosting [40] (histogram-based, maximum depth 3, learning rate 0.05, 300 iterations, minimum leaf size 50); change the exclusion threshold to 10% and 30%; raise transaction costs to 25 basis points; drop

How to cite this article:
Bhati N, Kataria A, Dhaundiyal P (November 06, 2026) Does Earnings-Call Climate Exposure Improve Machine-Learning Forecasts of Stock Downside Risk? Out-of-Sample Evidence from U.S. Firms. Cureus J Comput Sci 3 : es44389-026-00359-6.
DOI https://doi.org/10.7759/s44389-026-00359-6

formation year 2020; and use a more recent climate window ending in the first quarter of the formation year, which is valid only if the data use calendar quarters. To limit computing time, each model-based robustness run uses 20 linear placebos and one random-forest placebo.

Research transparency: corrections and exploratory analyses

The initial analysis revealed errors in the data construction and in two methodological choices. These issues were corrected before the results reported here were produced, and the outputs from the initial analysis have been retained. The following paragraphs describe the changes so that readers can assess their potential consequences.

The data errors concerned market values. In the initial analysis, prices of ADR listings were combined with SEC share counts that refer to ordinary shares, some cover-page share counts were reported on the wrong scale and share counts had not been adjusted for stock splits after the filing date. As a result, a small number of firms received implausibly large market values and dominated the value-weighted portfolios. The corrections, which exclude ADRs, require Form 10-K filers, adjust share counts for splits and screen out implausible share counts, are described in the section Identifier Matching and Market Values.

The methodological changes concerned specification and inference. The linear model originally used volatility in levels rather than logarithms, which prevented it from reproducing the naive log forecast and disadvantaged it relative to the random forest; volatility predictors now enter in logarithmic form. The initial analysis also relied on the firm bootstrap, which can understate uncertainty when forecast errors are correlated within a year. The analysis therefore places primary emphasis on the placebo permutation test, expands it to 100 and 10 draws and adds a comparison of the benchmark with SPY. Four additional analyses, not included in the original specification, were also conducted on the same sample: SHAP values for every test year, 100 instead of 10 random-forest placebos, a small-sample correction to the Diebold-Mariano test and equal-weighted portfolios. The earlier analysis with 10 random-forest placebos produced a permutation p-value of 0.182.

Results And Discussion

Experimental setup

All experiments use the 15,521 firm-years described in the Materials & Methods section, with test years from 2017 to 2024. The training set for each test year comprises all firm-years formed in earlier years, so it grows from 4,685 firm-years (formation years 2012-2016) for the 2017 forecasts to 13,821 firm-years (2012-2023) for the 2024 forecasts; the eight test sets together contain 10,836 firm-years. In each training window, the final year serves as the validation set for the random forest, whose settings are selected from six combinations, formed from a minimum leaf size of 20, 50 or 100 observations and a share of 0.33 or 0.66 of the predictors considered at each split, using 200 trees; the final forest has 300 trees. The naive forecast, models M1-M4 and the placebo models M2p and M4p (Table *4*) are evaluated on identical test observations, with 100 linear and 100 random-forest placebo draws.

Accuracy is summarized by the RMSE, the MAE and the out-of-sample R^2 against the naive forecast, both pooled over all test firm-years and by test year. Pairs of models are compared with a firm bootstrap (1,000 draws) and a Diebold-Mariano test on the eight annual loss differences, with and without the small-sample correction of Harvey et al. [41], while the permutation p-value in Equation (3) compares the real climate data with the placebos. The forecasts then form the portfolios in Table *5*, whose 96 monthly returns from July 2017 to June 2025 are evaluated after transaction costs of 10 basis points per side. All computations run in Python with the package versions listed under the Training and Validation section and a fixed random seed (42). The code, the package list and the saved outputs are available in the repository in Data Availability Statements section.

How to cite this article:
Bhati N, Kataria A, Dhaundiyal P (November 06, 2026) Does Earnings-Call Climate Exposure Improve Machine-Learning Forecasts of Stock Downside Risk? Out-of-Sample Evidence from U.S. Firms. Cureus J Comput Sci 3 : es44389-026-00359-6.
DOI https://doi.org/10.7759/s44389-026-00359-6

Descriptive statistics

Table 6 describes the estimation sample. Next-year downside semi-deviation has a mean of 28.9% and a median of 25.4%. Firms rarely discuss climate topics on earnings calls (Figure 2): overall exposure is zero in 11.6% of firm-years, opportunity exposure in 50.7%, regulatory exposure in 81.0% and physical exposure in 92.7%. In other words, most firms seldom raise regulatory or physical climate issues with analysts.

Variable	Mean	Median	SD	Share of zeros
Downside semi-deviation, next 12 months	0.2885	0.2541	0.1451	0.000
Volatility, next 12 months	0.4323	0.3723	0.2450	0.000
Maximum drawdown, next 12 months	−0.3365	−0.2970	0.1814	0.000
Log market capitalization	21.65	21.70	1.96	0.000
Overall climate exposure	0.0014	0.0004	0.0029	0.116
Regulatory exposure	0.0001	0.0000	0.0002	0.810
Physical exposure	0.0000	0.0000	0.0000	0.927
Opportunity exposure	0.0005	0.0000	0.0013	0.507

TABLE 6: Descriptive statistics (15,521 firm-years)

Risk measures are annualized decimals. Exposure measures are shares of transcript word pairs, as supplied in the data file.
SD, Standard Deviation

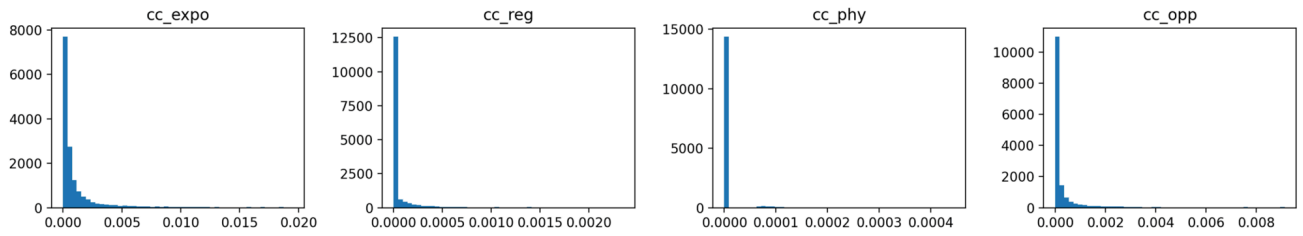


FIGURE 2: Distribution of climate exposure measures

Distribution of the four climate exposure measures across firm-years in the estimation sample. The panels show cc_expo (overall exposure), cc_reg (regulatory exposure), cc_phy (physical exposure) and cc_opp (opportunity exposure), measured as the share of transcript word pairs. The vertical axes show the number of firm-years.

How to cite this article:
Bhati N, Kataria A, Dhaundiyal P (November 06, 2026) Does Earnings-Call Climate Exposure Improve Machine-Learning Forecasts of Stock Downside Risk? Out-of-Sample Evidence from U.S. Firms. Cureus J Comput Sci 3 : es44389-026-00359-6.
DOI https://doi.org/10.7759/s44389-026-00359-6

In-sample association

Past risk dominates the panel regression (Table 7). Conditional on past risk, size, and year and industry effects, overall climate exposure has a small negative coefficient (–5.051, p = 0.015) and opportunity exposure a small positive one (9.005, p = 0.028). A one-standard-deviation difference in either measure changes predicted downside risk by only about 1.2-1.5%. Neither regulatory exposure (p = 0.086) nor physical exposure (p = 0.101) is significant at the 5% level. In the Fama-MacBeth regressions, no climate coefficient is significant, and the largest absolute t-statistic is 1.58. After controlling for past risk, the in-sample association between climate exposure and future downside risk is therefore weak and unstable.

Predictor	Coefficient	Standard error	p-value
Overall climate exposure	–5.051	2.070	0.015
Regulatory exposure	–18.917	11.019	0.086
Physical exposure	–65.358	39.908	0.101
Opportunity exposure	9.005	4.095	0.028
Log past volatility	0.538	0.028	<0.001
Log past downside semi-deviation	0.187	0.027	<0.001
Beta	–0.003	0.006	0.655
Log market capitalization	–0.026	0.002	<0.001
Book-to-market	0.002	0.005	0.752
Leverage	–0.013	0.008	0.096
Momentum	0.004	0.006	0.437
Observations / R^2	15,501 / 0.742		

TABLE 7: Panel regression of log next-year downside semi-deviation

Year and SIC-division fixed effects included; standard errors clustered by firm. The regression uses the 15,501 firm-years with complete predictor data.
SIC, Standard Industrial Classification

Out-of-sample forecast accuracy

Table 8 gives the accuracy of each model, and Table 9 compares the models in pairs; together they address RQ1 and RQ2. Adding climate exposure leaves forecast accuracy essentially where it was. The random forest's out-of-sample R^2 relative to the naive forecast rises from 10.25% to 10.30% (Table 8), and the MSE difference (–0.00008) has a bootstrap interval that contains zero and a Diebold-Mariano p-value of 0.665. In the linear model, climate exposure slightly increases the

How to cite this article:
Bhati N, Kataria A, Dhaundiyal P (November 06, 2026) Does Earnings-Call Climate Exposure Improve Machine-Learning Forecasts of Stock Downside Risk? Out-of-Sample Evidence from U.S. Firms. Cureus J Comput Sci 3 : es44389-026-00359-6.
DOI https://doi.org/10.7759/s44389-026-00359-6

MSE (+0.00002; p = 0.969). The random forest, by contrast, is more accurate than the linear model: it lowers the MSE by about 8% (−0.01053), the bootstrap interval excludes zero and the annual test gives p = 0.022. With only eight annual observations, however, the small-sample correction raises this p-value to 0.069, so the advantage is significant only at the 10% level (Table *9*).

Model	RMSE	MAE	Out-of-sample R^2 vs naive
Naive (last year's downside semi-deviation)	0.3682	0.2871	0.00%
M1: Linear	0.3636	0.2818	2.48%
M2: Linear + climate	0.3637	0.2818	2.47%
M3: Random forest	0.3488	0.2682	10.25%
M4: Random forest + climate	0.3487	0.2679	10.30%

TABLE 8: Out-of-sample forecast accuracy, test years 2017–2024 (10,836 firm-years)

RMSE and MAE are measured in units of the log target. Models M1–M4 are defined in Table 4.
RMSE, Root Mean Squared Error; MAE, Mean Absolute Error

Comparison	Difference in MSE	Firm-bootstrap 95% interval	Diebold–Mariano p-value
M2 vs M1 (climate, linear)	+0.00002	(−0.00027, 0.00031)	0.969
M4 vs M3 (climate, random forest)	−0.00008	(−0.00037, 0.00019)	0.665
M3 vs M1 (random forest vs linear)	−0.01053	(−0.01186, −0.00920)	0.022
M4 vs M2 (random forest vs linear, with climate)	−0.01062	(−0.01188, −0.00927)	0.015

TABLE 9: Pairwise comparisons of forecast accuracy

A negative difference means the first model is more accurate. Models M1–M4 are defined in Table 4.
MSE, Mean Squared Error

Most of the random forest's advantage over the naive forecast comes from a single period (Table *10*, exploratory; Figure *3*). In test year 2020 the forest removes 54.1% of the naive forecast's squared error. That year followed the COVID-19 shock: volatility fell back, and forecasts based on the previous year were far too high. In 2018, 2019 and 2021 the forest is

less accurate than the naive forecast. With test year 2020 left out, the pooled out-of-sample R^2 is −3.3% for the forest and −8.4% for the linear model. When all models are re-estimated without formation year 2020, the forest's R^2 is −1.1% (see Robustness), although it remains more accurate than the linear model.

Test year	Firm-years	M1 Linear	M3 Random forest	M4 Random forest + climate
2017	1,035	17.6%	28.7%	29.3%
2018	1,149	−5.2%	−0.6%	−0.3%
2019	1,207	−24.4%	−20.8%	−20.7%
2020	1,278	37.7%	54.1%	53.7%
2021	1,347	−20.8%	−7.0%	−6.9%
2022	1,468	−5.6%	1.5%	0.6%
2023	1,652	10.6%	9.1%	11.3%
2024	1,700	5.6%	12.3%	11.2%

TABLE 10: Out-of-sample R^2 against the naive forecast by test year (exploratory)

Models M1, M3 and M4 are defined in Table 4. Forecasts are made on 30 June of the test year for the following twelve months. These results were calculated after the main analysis from the saved out-of-sample predictions.

How to cite this article:
Bhati N, Kataria A, Dhaundiyal P (November 06, 2026) Does Earnings-Call Climate Exposure Improve Machine-Learning Forecasts of Stock Downside Risk? Out-of-Sample Evidence from U.S. Firms. Cureus J Comput Sci 3 : es44389-026-00359-6.
DOI https://doi.org/10.7759/s44389-026-00359-6

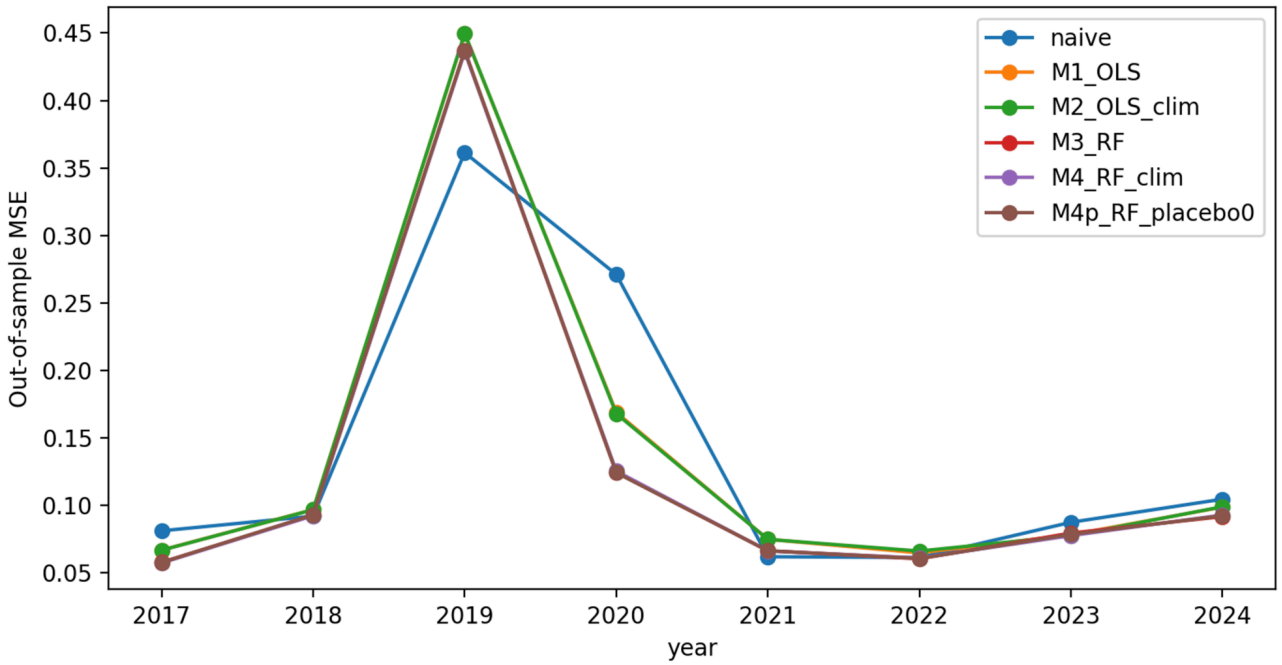


FIGURE 3: Out-of-sample forecast error by test year

Out-of-sample MSE by test year. Naive refers to last year's downside semi-deviation; M1_OLS and M2_OLS_clim refer to the linear models without and with climate exposure, respectively; M3_RF and M4_RF_clim refer to the random forest models without and with climate exposure, respectively; and M4p_RF_placebo0 refers to the random forest model with the first set of placebo climate variables.
MSE, Mean Squared Error; OLS, Ordinary Least Squares; RF, Random Forest

Placebo tests

For the random forest, the real climate data are indistinguishable from reassigned data: 16 of the 100 placebos perform at least as well as the real data (permutation p = 0.168; Table *11*). For the linear model, the real climate data outperform all 100 placebos (p = 0.010), yet the linear model without climate variables achieves an even lower MSE (0.132221, compared with 0.132239 with them). The climate variables therefore impair the linear forecast less than random numbers do, possibly because they correlate with existing predictors such as firm size, but they do not improve it. The placebo evidence does not alter the conclusion from the direct forecast comparisons.

Model	MSE, real climate data	Placebo mean (minimum, maximum)	Placebos at least as accurate	Permutation p-value
Linear (100 placebos)	0.132239	0.13241 (0.132240, 0.13253)	0	0.010
Random forest (10 placebos)	0.12161	0.12216 (0.12158, 0.12256)	1	0.182

How to cite this article:
Bhati N, Kataria A, Dhaundiyal P (November 06, 2026) Does Earnings-Call Climate Exposure Improve Machine-Learning Forecasts of Stock Downside Risk? Out-of-Sample Evidence from U.S. Firms. Cureus J Comput Sci 3 : es44389-026-00359-6.
DOI https://doi.org/10.7759/s44389-026-00359-6

TABLE 11: Real climate variables against placebo versions

Placebos reassign the four climate variables among firms in the same year and SIC division. The permutation p-value is defined in Equation (3).
MSE, Mean Squared Error; SIC, Standard Industrial Classification

Contributions of individual predictors

Averaged over the eight test years, past volatility makes the largest contribution to the predictions of the climate-informed random forest (mean absolute SHAP value 0.128; range across years 0.102-0.201), followed by past downside semi-deviation (0.084), firm size (0.065) and momentum (0.043) (Table *12*). The ranking is stable: past volatility ranks first in every year, and no climate variable ranks higher than seventh in any year. The four climate variables together contribute 0.011 on average (range 0.007-0.015), roughly a tenth of the contribution of past volatility, and physical exposure contributes almost nothing (0.0002). Past volatility and past downside semi-deviation are highly correlated ($r = 0.97$), so the SHAP values divide credit between them, and their individual values should be read jointly. The answer to RQ4 is that no climate dimension adds much to the forecasts, with physical exposure adding least.

Predictor	Mean absolute SHAP value
Log past volatility	0.1289
Log past downside semi-deviation	0.0858
Log market capitalization	0.0624
Momentum	0.0457
Beta	0.0176
Book-to-market	0.0166
Leverage	0.0062
Overall climate exposure	0.0052
Regulatory exposure	0.0047
Opportunity exposure	0.0034
Physical exposure	0.0001

TABLE 12: Mean absolute SHAP values for the random forest with climate data in the 2024 test year

Values are expressed in units of the log target. Industry indicators are omitted; each contributes 0.0026 or less.
SHAP, Shapley Additive Explanations.

How to cite this article:
Bhati N, Kataria A, Dhaundiyal P (November 06, 2026) Does Earnings-Call Climate Exposure Improve Machine-Learning Forecasts of Stock Downside Risk? Out-of-Sample Evidence from U.S. Firms. Cureus J Comput Sci 3 : es44389-026-00359-6.
DOI https://doi.org/10.7759/s44389-026-00359-6

Portfolio performance

The value-weighted benchmark tracks the U.S. market closely (Table *13*), which supports the reconstructed market values used as weights. Its largest holdings are Apple, Microsoft, Amazon and, in 2024, Nvidia, each making up 3-9% of the portfolio.

Portfolio	Annual return	Volatility	Maximum drawdown	Monthly ES (95%)
P0 benchmark	15.18%	17.09%	−24.78%	9.13%
SPY, same months	14.27%	16.44%	−23.93%	9.51%

TABLE 13: Benchmark validation, July 2017–June 2025

ES is calculated using monthly returns at the 95% confidence level.
ES, Expected Shortfall; SPY, SPDR S&P 500 ETF Trust

Table *14* presents the six portfolios, which address RQ3. Removing the stocks with the highest forecast risk reduces realized volatility and maximum drawdown only slightly relative to the benchmark. Adding climate information to the forecasts changes almost nothing: P3 and P2 have the same maximum drawdown (−24.21%), nearly identical monthly expected shortfall (8.96% and 8.97%) and tracking errors below 1% (Figures *4*-*5*). The regulatory-exposure screen (P1) earned 0.8 percentage points a year more than the benchmark, with slightly higher volatility. No Sharpe-ratio difference is statistically significant (Table *15*), including the difference between P1 and P0 ($p = 0.399$).

Portfolio	Return	Volatility	Sharpe	Sortino	Max. drawdown	Monthly ES (95%)	Tracking error	Turnover
P0 Benchmark	15.18%	17.09%	0.782	1.242	−24.78%	9.13%	—	0.039
P1 Regulatory screen	15.96%	17.39%	0.811	1.297	−24.74%	8.94%	1.70%	0.194
P2 RF, no climate	15.03%	16.86%	0.783	1.243	−24.21%	8.97%	0.70%	0.048
P3 RF + climate	15.11%	16.84%	0.787	1.251	−24.21%	8.96%	0.69%	0.048
P2L Linear, no climate	15.15%	17.01%	0.784	1.243	−24.70%	9.07%	0.24%	0.044
P3L Linear + climate	15.16%	17.00%	0.784	1.244	−24.71%	9.07%	0.23%	0.044

How to cite this article:
Bhati N, Kataria A, Dhaundiyal P (November 06, 2026) Does Earnings-Call Climate Exposure Improve Machine-Learning Forecasts of Stock Downside Risk? Out-of-Sample Evidence from U.S. Firms. Cureus J Comput Sci 3 : es44389-026-00359-6.
DOI https://doi.org/10.7759/s44389-026-00359-6

TABLE 14: Portfolio performance, July 2017–June 2025 (96 months; after transaction costs)

Portfolios P0–P3L are defined in Table 5. Returns and volatility are annualized. Sharpe and Sortino ratios use returns in excess of the risk-free rate. Transaction costs are 10 basis points per side. Turnover is the average one-way fraction of the portfolio traded at each annual rebalancing.
ES, Expected Shortfall; RF, Random Forest

Comparison	Difference in annualized Sharpe ratio	p-value
P3 vs P2 (climate, random forest forecasts)	+0.004	0.117
P3L vs P2L (climate, linear forecasts)	+0.000	0.489
P3 vs P0 benchmark	+0.005	0.652
P1 vs P0 benchmark (exploratory)	+0.029	0.399

TABLE 15: Tests of Sharpe-ratio differences

Tests follow Ledoit and Wolf [30] and use monthly returns in excess of the risk-free rate. Portfolios are defined in Table 5.

How to cite this article:
Bhati N, Kataria A, Dhaundiyal P (November 06, 2026) Does Earnings-Call Climate Exposure Improve Machine-Learning Forecasts of Stock Downside Risk? Out-of-Sample Evidence from U.S. Firms. Cureus J Comput Sci 3 : es44389-026-00359-6.
DOI https://doi.org/10.7759/s44389-026-00359-6

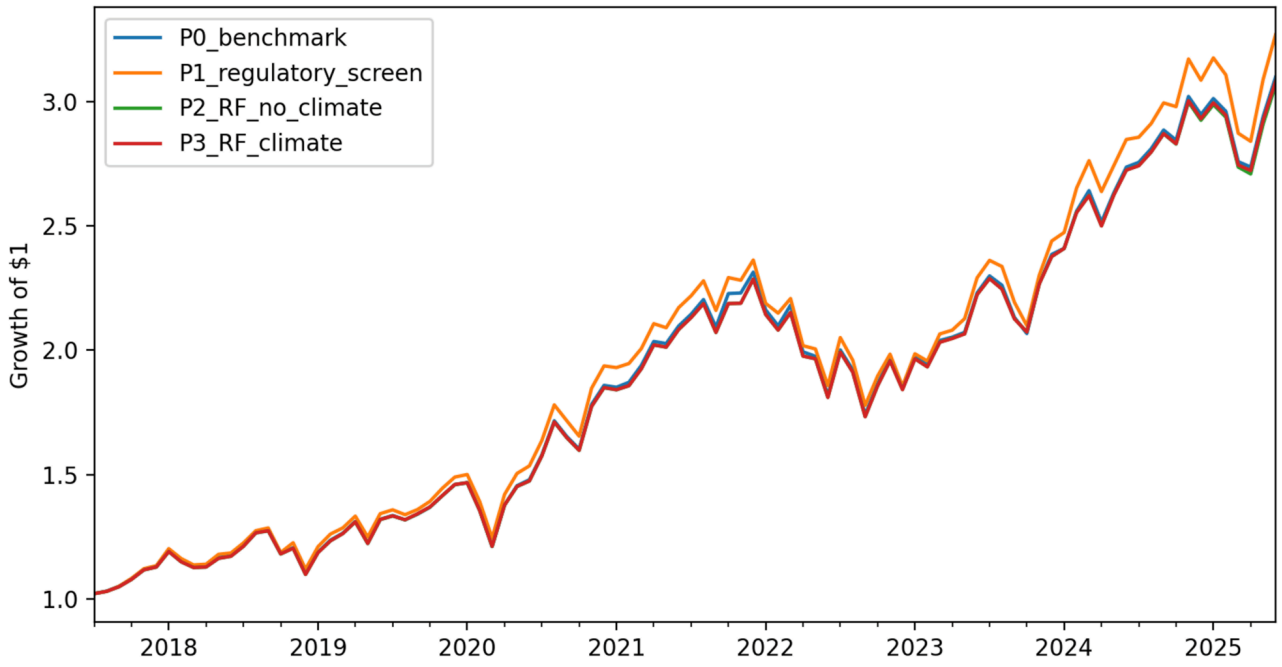


FIGURE 4: Cumulative portfolio returns

Cumulative value of $1 invested in portfolios P0–P3, July 2017–June 2025, after transaction costs. P0 denotes value-weighted benchmark; P1 denotes regulatory-exposure screen; P2 and P3 denote random-forest forecasts without and with climate data, respectively. The lines for P0, P2 and P3 largely overlap.

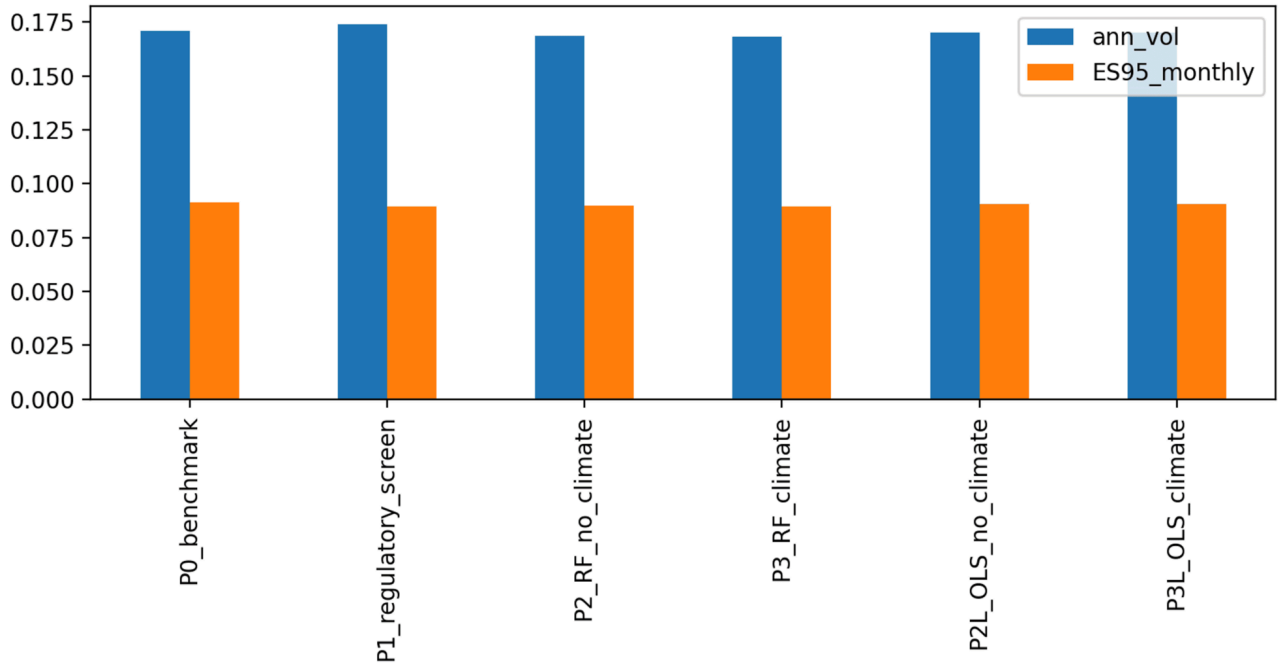


FIGURE 5: Portfolio risk comparison

Annualized volatility (ann_vol) and monthly expected shortfall at the 95% level (ES95_monthly) of the six portfolios, July 2017–June 2025.

How to cite this article:
Bhati N, Kataria A, Dhaundiyal P (November 06, 2026) Does Earnings-Call Climate Exposure Improve Machine-Learning Forecasts of Stock Downside Risk? Out-of-Sample Evidence from U.S. Firms. Cureus J Comput Sci 3 : es44389-026-00359-6.
DOI https://doi.org/10.7759/s44389-026-00359-6

Robustness

Table *16* repeats the main comparisons under each robustness check. Climate information never improves the tree-based forecasts, and with the more recent climate window it makes them significantly worse (+0.0012). The tree-based models outperform the linear models in every variant, including the one in which gradient boosting replaces the random forest. The expected shortfall of portfolios with and without climate information differs by at most 0.09 percentage points. Because each robustness run uses a single random-forest placebo, too few for a permutation test, we report only the linear permutation p-values. With equal-weighted instead of value-weighted portfolios, climate information again leaves risk essentially unchanged: P3 and P2 have annualized volatility of 20.24% and 20.25% and monthly expected shortfall of 11.73% and 11.77%. The Sharpe ratio of P3 is slightly higher than that of P2 (0.524 against 0.514), and this small difference is statistically significant (p = 0.016). All equal-weighted screened portfolios have lower Sharpe ratios than the equal-weighted benchmark (P3 against P0: 0.524 against 0.630, p = 0.018). These equal-weighted results are available in the repository.

Variant	Climate effect, tree model (M4 vs M3)	Tree vs linear (M3 vs M1)	Linear permutation p (20 placebos)	Monthly ES, P2 / P3	Sharpe, P2 / P3
Main specification	−0.0001 (−0.0004, 0.0002)	−0.0105 (−0.0119, −0.0092)	0.010 (100 placebos)	8.97% / 8.96%	0.783 / 0.787
Volatility as target	+0.0001 (−0.0002, 0.0003)	−0.0082 (−0.0094, −0.0069)	0.095	8.97% / 8.97%	0.790 / 0.790
Formation year 2020 removed	−0.0002 (−0.0004, 0.0001)	−0.0043 (−0.0055, −0.0030)	0.048	8.95% / 9.04%	0.598 / 0.593
Gradient boosting	+0.0001 (−0.0003, 0.0004)	−0.0113 (−0.0128, −0.0099)	0.048	8.96% / 8.98%	0.751 / 0.743
More recent climate window	+0.0012 (0.0009, 0.0016)	−0.0105 (−0.0119, −0.0092)	0.048	9.03% / 9.11%	0.786 / 0.787
10% exclusion	—	—	—	9.10% / 9.10%	0.785 / 0.784
30% exclusion	—	—	—	8.89% / 8.89%	0.758 / 0.757
Costs of 25 basis points	—	—	—	8.97% / 8.96%	0.781 / 0.786

TABLE 16: Robustness checks

Models M1–M4 are defined in Table 4, and portfolios P2–P3 are defined in Table 5. Differences in MSE are reported with firm-bootstrap 95% confidence intervals in parentheses; a negative value indicates that the first model is more accurate. Figures are rounded to four decimal places, so the main-specification row matches Table 9 after rounding. In the gradient-boosting row, M3 and M4 denote gradient-boosting models. Removing formation year 2020 leaves 84 portfolio months; all other rows use 96 months. Dashes indicate portfolio-only variants for which the forecasts are unchanged.
ES, Expected Shortfall; MSE, Mean Squared Error

How to cite this article:
Bhati N, Kataria A, Dhaundiyal P (November 06, 2026) Does Earnings-Call Climate Exposure Improve Machine-Learning Forecasts of Stock Downside Risk? Out-of-Sample Evidence from U.S. Firms. Cureus J Comput Sci 3 : es44389-026-00359-6.
DOI https://doi.org/10.7759/s44389-026-00359-6

Summary of hypothesis tests

Table *17* summarizes how each hypothesis fared. Only H2 receives support, and only partly: the random forest is more accurate than the linear model, but the advantage is marginally significant once the small-sample correction is applied. Climate exposure improves neither the forecasts nor the risk of the portfolios.

Hypothesis	Outcome	Main evidence
H1: Climate exposure improves out-of-sample forecasts	Not supported	Tables 9, 11, 16
H1b: Regulatory exposure contributes more than physical exposure	Not supported as stated: no dimension contributes meaningfully, although physical exposure contributes least	Tables 7, 12
H2: Random forest and linear model differ in accuracy	Supported: the random forest is more accurate	Tables 9, 16
H3: Climate-informed portfolios have lower downside risk	Not supported	Tables 14, 15

TABLE 17: Hypotheses and outcomes

Discussion

What the Results Show

The central result is the absence of a predictive gain. In this sample, climate exposure measured at the firm level from earnings calls does not materially improve one-year-ahead forecasts of downside risk once past risk and standard firm characteristics are known. This holds for linear and nonlinear models, for both risk measures and across the robustness checks. The random forest using the real climate data performs no better than versions in which those data are shuffled. The portfolio tests point the same way: forecasts with and without climate data lead to portfolios with nearly identical risk.

What the Results Do Not Show

The results do not show that climate risk is irrelevant for finance or that climate exposure has no bearing on asset prices. The design tests how accurately risk can be forecast, not how it is priced, and it is compatible with evidence that climate exposure is associated with volatility [6] and reflected in option and equity prices [4,5]. Nor do the results show that climate data are of no use for other investment purposes, such as reporting, engagement or matching portfolios to investors' preferences. The findings concern one measure, one market, one horizon and one period. Because the design is predictive, it cannot establish whether climate exposure causes changes in risk. A non-significant improvement is also not proof that climate information has no predictive value. The findings are specific to earnings-call exposure measures, U.S.-listed non-financial firms, a one-year horizon and the 2017-2024 test period. Other climate indicators, horizons or market conditions may produce different results.

Possible Explanations

How to cite this article:
Bhati N, Kataria A, Dhaundiyal P (November 06, 2026) Does Earnings-Call Climate Exposure Improve Machine-Learning Forecasts of Stock Downside Risk? Out-of-Sample Evidence from U.S. Firms. Cureus J Comput Sci 3 : es44389-026-00359-6.
DOI https://doi.org/10.7759/s44389-026-00359-6

Several explanations fit the null result. Our design cannot tell them apart, so we present them as possibilities rather than established mechanisms. Volatility is highly persistent, and in every model, past volatility and past downside semi-deviation carry most of the predictive content (Table *7* and Table *12*), so climate-related risk that has already affected a stock may be absorbed by these variables. If climate exposure is already reflected in prices and recent volatility, as the pricing evidence suggests, little is left for it to contribute to a forecast. The measures are also sparse because most firms rarely raise regulatory or physical climate topics on earnings calls (Table *6*), and they change slowly relative to a one-year horizon. The conservative window adds to this lag: under the calendar reading, the calls it averages took place between about 9 and 21 months before formation, and under the fiscal reading, possibly earlier. The age of the measure does not, however, appear to explain the null result, because a fresher window ending in the first quarter of the formation year makes the forecasts slightly worse rather than better (Table *16*). Finally, climate-related risk may matter over longer horizons or in rare and severe events that a 12-month downside-deviation measure averages out.

The in-sample coefficients reported here differ in sign from the positive association documented by Naseer et al. [6]. That study differs from ours in specification, control variables, countries and period, and our coefficients are small and unstable across estimation methods (Table *7*). These estimates should therefore not be read as evidence against theirs.

What Machine Learning Adds

The results separate the value of the model from the value of the climate data. The random forest's advantage over the linear model appears in every variant, although, with eight test years, it is only marginally significant after a small-sample correction, consistent with evidence that tree-based methods capture nonlinear structure in financial data [8]. Its advantage over the naive forecast is less secure. It arises mainly in the year in which volatility fell sharply after the COVID-19 shock, most likely because a model trained on many firms and years captures the tendency of volatility to revert after a spike, which last year's value cannot. Outside such periods, the naive forecast proves difficult to outperform, in line with evidence that machine-learning gains in volatility forecasting can be modest [9]. Pooled accuracy statistics can therefore conceal a dependence on a few turbulent periods, which is why we report accuracy by year.

This distinction clarifies what the term "AI-powered" in the title amounts to here. AI-powered climate risk assessment pairs a flexible model with climate data. In this study, the flexible model improves on the linear model, but the climate data improve the forecasts of neither model.

Interpreting the Portfolio Results

Because forecasts with and without climate data hardly differ, the portfolios built from them hardly differ either, and the small reductions in volatility and drawdown relative to the benchmark come from the conventional predictors. The higher realized return of the regulatory screen should be interpreted with caution: its Sharpe ratio cannot be statistically distinguished from the benchmark's, and eight years is a short span over which to judge a screening rule. With equal weights, the climate-informed random-forest screen has a slightly but significantly higher Sharpe ratio than the screen without climate information (0.524 against 0.514), while its risk is essentially the same. Given the number of portfolio comparisons and the small size of the difference, we treat this as a result that requires confirmation rather than as evidence of economic value. None of the portfolio results is intended as an investment recommendation.

Implications

For investors and portfolio managers, the results argue against treating earnings-call climate exposure scores as signals of short-horizon risk, since in this sample they add nothing to forecasts based on past risk and basic firm characteristics. Climate screens may serve other purposes, but this evidence does not support using them to reduce one-year downside risk. For financial institutions and supervisors, the results suggest that text-based exposure measures should not be treated as interchangeable with forward-looking risk indicators and that out-of-sample evidence is needed before such scores are used to predict risk. Because the study is about how investors forecast stock risk, it has no direct implication for how firms should manage or disclose their climate exposure. For researchers, the results support three practices:

How to cite this article:
Bhati N, Kataria A, Dhaundiyal P (November 06, 2026) Does Earnings-Call Climate Exposure Improve Machine-Learning Forecasts of Stock Downside Risk? Out-of-Sample Evidence from U.S. Firms. Cureus J Comput Sci 3 : es44389-026-00359-6.
DOI https://doi.org/10.7759/s44389-026-00359-6

checking in-sample associations in out-of-sample tests, using placebo controls to separate the value of the data from side effects of the model, and reporting forecast gains period by period, because a single turbulent year can create an apparent advantage for machine learning.

Limitations

Survivorship bias is the most important limitation. Because we matched firms through lists of current tickers, we could include only 2,490 of the 6,279 U.S. firms with climate data, and most of the remainder appear to have been delisted or acquired. The problem is larger in earlier years: the share of U.S. firms with climate data that could be matched rises from 37.8% in formation year 2012 and 47.7% in 2017, the first test year, to 65.4% in 2024. Because the sample is conditioned on survival to the present, its composition reflects information that was not available at the time, although no predictor uses future information. Surviving firms are likely to be less risky than those that disappeared, which may understate risk levels; comparisons between models are less affected because all models use the same firms. Delisting returns are not freely available, so we treat a stock that stops trading as cash for the rest of the year. This assumption sets the return after delisting to zero, whereas delistings for poor performance are often preceded or followed by large losses. It is therefore likely to understate the downside risk of all portfolios, including the benchmark, by a similar amount, so its effect on comparisons between screened portfolios should be limited.

Market values had to be reconstructed from prices and SEC share counts. They are missing in 567 firm-years with implausible share counts; 1,855 firm-years rely on weighted-average share counts; and in 5,348 firm-years, total liabilities equal assets minus equity. The climate measure covers only firms that hold earnings calls. Its keyword list was learned from the full sample period, and its documentation does not state whether the quarter field refers to calendar or fiscal quarters, although the conservative climate window used here is valid under either reading. Regulatory and physical exposure are zero in most firm-years, which limits what the data can reveal about these dimensions.

The sample covers U.S.-listed non-financial firms, with portfolios formed from 2012 to 2024, returns measured to June 2025, and a one-year forecast horizon. Results may differ for other markets, longer horizons or other climate measures such as emissions. With eight test years, time-series tests have limited power, and the random forest's advantage over the linear model is only marginally significant once a small-sample correction is applied. Finally, the design identifies predictive content rather than causal effects.

Future Research

These limitations suggest several extensions. Repeating the analysis with price data that include delisted firms would show whether survivorship bias affects the conclusions. Longer forecast horizons or outcomes focused on rare but severe losses might reveal climate effects that a one-year downside-deviation measure misses. Firm-level emissions, disclosure-based measures taken from annual reports, or physical-risk scores based on where firms' assets are located would show whether the null result is specific to earnings-call language. The design could also be applied to other asset classes, such as corporate and green bonds, where climate information may bear on credit and pricing risk. Applying the design outside the United States, where climate policy and disclosure differ, would test how far the findings generalize. More random-forest placebos, together with a forecast evaluation that allows for correlation across firms within a year, would also sharpen the statistical tests.

Conclusions

This study examined whether firm-level climate change exposure measured from earnings calls helps to forecast the downside risk of U.S. stocks and to improve risk-screened portfolios, and whether a machine-learning model adds value to this task. Climate exposure does not materially improve out-of-sample forecasts of next-year downside risk: accuracy remains essentially unchanged when climate exposure is added to linear models, random forests or gradient boosting, and the random forest with real climate data cannot be distinguished from its placebo versions. Random forests forecast downside risk more accurately than linear models, but with eight test years the advantage is only marginally significant,

How to cite this article:
Bhati N, Kataria A, Dhaundiyal P (November 06, 2026) Does Earnings-Call Climate Exposure Improve Machine-Learning Forecasts of Stock Downside Risk? Out-of-Sample Evidence from U.S. Firms. Cureus J Comput Sci 3 : es44389-026-00359-6.
DOI https://doi.org/10.7759/s44389-026-00359-6

and the gain of the random forest over a simple past-risk forecast is sensitive to the recovery year after the COVID-19 shock. Adding climate information leaves the risk of the portfolios almost unchanged. No climate dimension contributes meaningfully to the forecasts, and physical exposure contributes least.

For U.S.-listed firms at a one-year horizon, the association between climate exposure and stock risk reported in earlier work does not translate into a forecasting advantage once past risk is known. For research that combines machine learning and climate risk, the methodological implication is that the contributions of climate data and model flexibility should be evaluated separately and through out-of-sample tests.

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Concept and design: Nandini Bhati, Abhinav Kataria , Pankaj Dhaundiyal

Acquisition, analysis, or interpretation of data: Nandini Bhati

Drafting of the manuscript: Nandini Bhati, Abhinav Kataria , Pankaj Dhaundiyal

Critical review of the manuscript for important intellectual content: Nandini Bhati

Supervision: Abhinav Kataria , Pankaj Dhaundiyal

Disclosures

Human subjects: All authors have confirmed that this study did not involve human participants or tissue.

Animal subjects: All authors have confirmed that this study did not involve animal subjects or tissue.

Conflicts of interest: In compliance with the ICMJE uniform disclosure form, all authors declare the following:

- **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work.
- **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work.
- **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

Data Availability Statements

The code, package versions, constructed firm-year data, forecasts and portfolio returns are openly available in Zenodo at https://doi.org/10.5281/zenodo.23056613. Climate exposure data are available from OSF (https://doi.org/10.17605/OSF.IO/FD6JQ); accounting data from SEC EDGAR, identifiers from OpenFIGI, prices from Yahoo Finance and the risk-free rate from the Kenneth R. French Data Library. Raw prices are not redistributed under the provider's terms.

Acknowledgements

Use of generative AI tools: An AI assistant (Claude, Anthropic) was used to support literature searches and source verification, to write and debug the analysis code, and to draft and edit the manuscript text. The analyses were run by the first author on the data described in Materials & Methods section; all authors reviewed the content and take full responsibility for the manuscript.

References

1. Network for Greening the Financial System: A call for action: climate change as a source of financial risk. First comprehensive report. (2019). Accessed: September 30, 2026: https://www.ngfs.net/en/publications-and-statistics/publications/first-comprehensive-report-call-action.
2. Giglio S, Kelly B, Stroebel J: Climate finance. Annual Review of Financial Economics. 2021, 13:15-36. 10.1146/annurev-financial-102620-103311
3. Acharya VV, Berner R, Engle R, Jung H, Stroebel J, Zeng X, Zhao Y: Climate stress testing. Annual Review of Financial Economics. 2023, 15:291-326. 10.1146/annurev-financial-110921-101555
4. Ilhan E, Sautner Z, Vilkov G: Carbon tail risk. The Review of Financial Studies. 2021, 34:1540-1571. 10.1093/rfs/hhaa071
5. Sautner Z, van Lent L, Vilkov G, Zhang R: Firm-level climate change exposure. The Journal of Finance. 2023, 78:1449-1498. 10.1111/jofi.13219
6. Naseer MM, Guo Y, Zhu X: The dynamics of corporate climate risk and market volatility: International evidence. The North American Journal of Economics and Finance. 2025, 78:102435. 10.1016/j.najef.2025.102435
7. Rolnick D, Donti PL, Kaack LH, et al.: Tackling climate change with machine learning. ACM Computing Surveys. 2022, 55:1-96. 10.1145/3485128
8. Gu S, Kelly B, Xiu D: Empirical asset pricing via machine learning. The Review of Financial Studies. 2020, 33:2223-2273. 10.1093/rfs/hhaa009
9. Christensen K, Siggaard M, Veliyev B: A machine learning approach to volatility forecasting. Journal of Financial Econometrics. 2023, 21:1680-1727. 10.1093/jjfinec/nbac020
10. Awijen H, Ben Jabeur S, Pillot J: Interpretable machine learning models for ESG stock prices under transition and physical climate risk. Annals of Operations Research. 2026, 357:191-221. 10.1007/s10479-024-06231-x
11. Bonato M, Cepni O, Gupta R, Pierdzioch C: Climate risks and state-level stock market realized volatility. Journal of Financial Markets. 2023, 66:100854. 10.1016/j.finmar.2023.100854
12. Bouyaddou Y, Jebabli I: Integration of investor behavioral perspective and climate change in reinforcement learning for portfolio optimization. Research in International Business and Finance. 2025, 73:102639. 10.1016/j.ribaf.2024.102639
13. Zhou M, Ma Y: Climate risk and predictability of global stock market volatility. Journal of International Financial Markets, Institutions and Money. 2025, 101:102135. 10.1016/j.intfin.2025.102135
14. Bolton P, Kacperczyk M: Do investors care about carbon risk?. Journal of Financial Economics. 2021, 142:517-549. 10.1016/j.jfineco.2021.05.008
15. Pástor Ľ, Stambaugh RF, Taylor LA: Sustainable investing in equilibrium. Journal of Financial Economics. 2021, 142:550-571. 10.1016/j.jfineco.2020.12.011
16. Faccini R, Matin R, Skiadopoulos G: Dissecting climate risks: Are they reflected in stock prices?. Journal of Banking & Finance. 2023, 155:106948. 10.1016/j.jbankfin.2023.106948
17. Hong H, Li FW, Xu J: Climate risks and market efficiency. Journal of Econometrics. 2019, 208:265-281. 10.1016/j.jeconom.2018.09.015
18. Kölbel JF, Leippold M, Rillaerts J, Wang Q: Ask BERT: how regulatory disclosure of transition and physical climate risks affects the CDS term structure. Journal of Financial Econometrics. 2024, 22:30-69. 10.1093/jjfinec/nbac027
19. Bingler JA, Kraus M, Leippold M, Webersinke N: Cheap talk and cherry-picking: What ClimateBert has to say on corporate climate risk disclosures. Finance Research Letters. 2022, 47:102776. 10.1016/j.frl.2022.102776
20. Nguyen Q, Diaz-Rainey I, Kuruppuarachchi D: Predicting corporate carbon footprints for climate finance risk analyses: A machine learning approach. Energy Economics. 2021, 95:105129. 10.1016/j.eneco.2021.105129
21. Salisu AA, Pierdzioch C, Gupta R, van Eyden R: Climate risks and U.S. stock-market tail risks: a forecasting experiment using over a century of data. International Review of Finance. 2023, 23:228-244. 10.1111/irfi.12397
22. Breiman L: Random forests. Machine Learning. 2001, 45:5-32. 10.1023/a:1010933404324

How to cite this article:
Bhati N, Kataria A, Dhaundiyal P (November 06, 2026) Does Earnings-Call Climate Exposure Improve Machine-Learning Forecasts of Stock Downside Risk? Out-of-Sample Evidence from U.S. Firms. Cureus J Comput Sci 3 : es44389-026-00359-6.
DOI https://doi.org/10.7759/s44389-026-00359-6

23. Assael J, Carlier L, Challet D: Dissecting the explanatory power of ESG features on equity returns by sector, capitalization, and year with interpretable machine learning. Journal of Risk and Financial Management. 2023, 16:159. 10.3390/jrfm16030159

24. Cerqueira V, Torgo L, Mozetič I: Evaluating time series forecasting models: an empirical study on performance estimation methods. Machine Learning. 2020, 109:1997-2028. 10.1007/s10994-020-05910-7

25. Lundberg SM, Lee SI: A unified approach to interpreting model predictions [PREPRINT]. arXiv. 2017, 10.48550/arXiv.1705.07874

26. Lundberg SM, Erion G, Chen H, et al.: From local explanations to global understanding with explainable AI for trees. Nature Machine Intelligence. 2020, 2:56-67. 10.1038/s42256-019-0138-9

27. Engle RF, Giglio S, Kelly B, Lee H, Stroebel J: Hedging climate change news. The Review of Financial Studies. 2020, 33:1184-1216. 10.1093/rfs/hhz072

28. Sautner Z, van Lent L, Vilkov G, Zhang R: Firm-level climate change exposure. OSF. (2025). Accessed: November 5, 2026: https://doi.org/10.17605/OSF.IO/FD6JQ.

29. U.S. Securities and Exchange Commission: EDGAR application programming interfaces (APIs). (2025). Accessed: September 30, 2026: https://www.sec.gov/search-filings/edgar-application-programming-interfaces.

30. OpenFIGI: API documentation. Accessed: September 30, 2026: https://www.openfigi.com/api/documentation.

31. Aroussi R: yfinance: download market data from Yahoo! Finance's API. Accessed: September 30, 2026: https://github.com/ranaroussi/yfinance.

32. French KR: Data library. Tuck School of Business, Dartmouth College. Accessed: September 25, 2026: https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

33. Probst P, Wright MN, Boulesteix AL: Hyperparameters and tuning strategies for random forest. WIREs Data Mining and Knowledge Discovery. 2019, 9:e1301. 10.1002/widm.1301

34. Pedregosa F, Varoquaux G, Gramfort A, et al.: Scikit-learn: machine learning in Python. Journal of Machine Learning Research. 2011, 12:2825-2830.

35. Seabold S, Perktold J: Statsmodels: econometric and statistical modeling with Python. Proceedings of the 9th Python in Science Conference. 2010, 92-96. 10.25080/Majora-92bf1922-011

36. Bhati N, Kataria A, Dhaundiyal P: Code and data for "Does Earnings-Call Climate Exposure Improve Machine-Learning Forecasts of Stock Downside Risk? Out-of-Sample Evidence from U.S. Firms" [Computer software]. Zenodo. (2026). Accessed: November 5, 2026: https://doi.org/10.5281/zenodo.23229046.

37. Diebold FX, Mariano RS: Comparing predictive accuracy. Journal of Business & Economic Statistics. 1995, 13:253-263. 10.2307/1392185

38. Fama EF, MacBeth JD: Risk, return, and equilibrium: empirical tests. Journal of Political Economy. 1973, 81:607-636. 10.1086/260061

39. Ledoit O, Wolf M: Robust performance hypothesis testing with the Sharpe ratio. Journal of Empirical Finance. 2008, 15:850-859. 10.1016/j.jempfin.2008.03.002

40. Friedman JH: Greedy function approximation: a gradient boosting machine. The Annals of Statistics. 2001, 29:1189-1232. 10.1214/aos/1013203451

41. Harvey D, Leybourne S, Newbold P: Testing the equality of prediction mean squared errors. International Journal of Forecasting. 1997, 13:281-291. 10.1016/s0169-2070(96)00719-4

How to cite this article:
Bhati N, Kataria A, Dhaundiyal P (November 06, 2026) Does Earnings-Call Climate Exposure Improve Machine-Learning Forecasts of Stock Downside Risk? Out-of-Sample Evidence from U.S. Firms. Cureus J Comput Sci 3 : es44389-026-00359-6.
DOI https://doi.org/10.7759/s44389-026-00359-6