

Climate Risk, Firm Resilience, and Supply-Chain Network Propagation in the S&P 500

Kian Jindal

Heritage International Xperiential School Pune

DOI: 10.46609/IJSSER.2026.v11i09.033 URL: <https://doi.org/10.46609/IJSSER.2026.v11i09.033>

Received: 8 September 2026 / Accepted: 20 September 2026 / Published: 28 September 2026

ABSTRACT

Firm-level climate assessments can miss stress transmitted through material customers. This study asks whether incorporating publicly disclosed customer dependence changes scenario-based estimates of financial resilience for 503 S&P 500 firms. SEC financial statements and annual 10-K disclosures are combined with a study-authored, NGFS-inspired climate calibration for Current Policies and Net Zero 2050 in 2030, 2040, and 2050. Baseline resilience is measured with a continuous accounting-based Altman Z" proxy. Direct sector shocks are applied to selected accounting variables, and customer stress is transmitted upstream through 612 cleaned customer-to-supplier links. Under Current Policies, mean direct stress rises from 2.10 percent in 2030 to 10.37 percent in 2050; the principal disclosed-network contribution rises from 1.33 to 7.04 percentage points. Net Zero 2050 front-loads adjustment and produces lower modeled stress by 2050. However, a strict-extraction sensitivity retains only 27 edges and greatly reduces average network exposure, showing that conclusions depend on disclosure and parsing quality. A residual graph model also fails to outperform self-only and shuffled controls. The study therefore contributes an auditable screening framework, not a forecast of losses, default, or bankruptcy.

Keywords: Climate risk; supply-chain networks; financial resilience; climate stress testing; customer concentration; S&P 500

1. Introduction

Climate change can affect corporate financial resilience through physical hazards and the transition toward lower-carbon production. Acute events can damage assets and interrupt transport, while chronic heat, water scarcity, and sea-level change can reduce productivity or raise operating costs. Carbon prices, regulation, technological change, litigation, and shifting demand can also alter revenue, margins, and asset values. These effects matter beyond the

directly exposed firm because customers and suppliers connect production and cash flow across sectors (Dietz et al., 2016; Hong, Karolyi, and Scheinkman, 2020; NGFS, 2024a).

A firm's own sector is therefore only one part of its exposure. A supplier with modest direct climate sensitivity may depend on a customer in energy, utilities, manufacturing, or another stressed activity. If that customer reduces orders, renegotiates contracts, or delays payments, stress can move upstream. Research on disasters and production networks shows that shocks can travel through buyer-supplier relationships and persist beyond the initially affected company (Barrot and Sauvagnat, 2016; Carvalho et al., 2021). Yet practical climate screens often combine firm accounts with sector shocks while treating firms as independent observations.

The central research question is: To what extent does incorporating publicly disclosed customer dependencies change scenario-based estimates of financial resilience among S&P 500 firms from 2030 to 2050?

The study pursues six objectives:

1. Measure baseline accounting-based resilience across the 503-firm analytical sample.
2. Compare direct climate stress under Current Policies and Net Zero 2050 in 2030, 2040, and 2050.
3. Estimate the additional upstream exposure associated with disclosed material-customer relationships.
4. Determine whether network information changes firm and sector vulnerability rankings.
5. Identify firms with low modeled direct exposure but elevated customer-linked exposure.
6. Test sensitivity to financial proxies, extraction quality, fallback edge weights, shock magnitude, and hidden-vulnerability thresholds.

The analysis combines source-first public financial statements, major-customer disclosures, a constructed climate calibration, and a directed customer-to-supplier graph. The accounting outcome is a continuous Altman Z" resilience proxy rather than a probability of default. The principal network method is a transparent weighted exposure calculation. A graph neural network (GNN) is retained only as a secondary diagnostic because the observed graph does not outperform self-only and shuffled-edge controls.

The study's contribution is deliberately narrow. It provides a reproducible way to ask whether disclosed customer dependence changes comparative climate-risk screens and makes the evidential status of every input explicit. Financial statements and filing excerpts are observed

data; financial fallbacks and relationship confidence weights are supported substitutions; sector mappings and climate shocks are researcher-constructed inputs; and all scenario results are conditional model outputs. This separation is necessary because a precise-looking scenario number is not an observed future loss.

The expected pattern is temporal and relational rather than causal. Current Policies should become more stressful over the long horizon because the calibration assigns increasing physical damage, while Net Zero 2050 should impose more adjustment near 2030 and less physical stress later. Among firms with observed customer links, suppliers connected to more highly stressed customers should receive larger modeled network exposure. Whether observed topology improves a learned accounting model remains an open diagnostic question.

The study also treats inconvenient evidence as part of its contribution. The broad disclosed network creates much larger estimated exposure than the strict-pattern subset, and the observed graph does not outperform basic machine-learning controls. Reporting these results prevents coverage gained through automated extraction from being mistaken for measurement certainty and identifies where better disclosure or manual validation would improve the analysis.

The remainder of the paper separates prior research from the study design. Section 2 reviews corporate climate risk, scenario analysis, financial resilience, production networks, and graph methods and then states the research gap. Section 3 explains the methodology, including shock construction, accounting transformations, relationship extraction, edge weighting, and robustness tests. Section 4 describes the data. Section 5 presents and interprets the results. Section 6 discusses practical implications, and Section 7 gives a concise answer to the research question.

2. Literature Review

2.1 Corporate Climate Risk and Scenario Analysis

Corporate climate risk is commonly divided into physical and transition components. Physical risk includes acute events such as floods, storms, wildfires, and heat waves and chronic changes such as rising temperatures, altered rainfall, water scarcity, and sea-level rise. These hazards can impair assets, interrupt work, raise insurance costs, reduce labour productivity, and disrupt logistics. Transition risk arises from policy, technology, legal, market, and reputational changes associated with decarbonisation. Carbon prices, emissions standards, renewable technologies, energy-demand changes, and investor preferences can affect operating margins and strand assets (TCFD, 2017; NGFS, 2024a).

Empirical finance research increasingly finds that climate exposure is reflected in corporate outcomes, although measurement remains contested. Bolton and Kacperczyk (2021) report a

carbon premium in U.S. equity returns, while Li et al. (2024) show that climate-risk measures extracted from earnings calls contain information about firms' exposures and responses. Aswani, Raghunandan, and Rajgopal (2024) demonstrate that conclusions about emissions and returns depend on whether emissions are disclosed or estimated. This literature supports examining financial consequences but warns against treating constructed exposure measures as observed losses.

Direct exposure and financial consequence are not identical. A carbon-intensive firm may face substantial transition pressure yet remain financially resilient because it has strong margins, liquidity, or access to capital. Conversely, a company with modest operational emissions may be fragile because it is highly leveraged or depends on a narrow customer base. Climate-risk assessment therefore needs both an external stress indicator and a representation of the firm's capacity to absorb it.

Scenario analysis addresses the lack of historical observations for long-horizon climate pathways. NGFS scenarios provide coherent narratives that combine transition and physical risk rather than probability-weighted forecasts (NGFS, 2024a, 2024b). Current Policies generally places less adjustment in the near term but allows greater long-run warming and physical damage. Net Zero 2050 requires stronger early transition but limits later physical risk. Central-bank exercises use such pathways to compare vulnerabilities under common assumptions (Alogoskoufis et al., 2021; Bank of England, 2022). Translating macro pathways into firm-level financial shocks nevertheless requires additional modeling choices, especially when geographic assets, emissions, adaptation, and detailed sector production functions are unavailable (Acharya et al., 2023).

That translation creates two common risks. Sector averages can conceal substantial variation in technology, geography, regulation, and business mix. A model can also mix sourced scenario variables with researcher assumptions so thoroughly that readers cannot identify which values came from an external dataset. Good practice is therefore to publish the mapping, parameters, formulas, and sensitivity ranges and to label the resulting figures as conditional estimates.

2.2 Accounting-Based Financial Resilience

Liquidity, accumulated earnings, profitability, and leverage affect a firm's ability to absorb stress. Altman's original Z-score combined accounting and market ratios to discriminate between failed and non-failed manufacturers (Altman, 1968). Later versions adapted the specification for private and non-manufacturing firms (Altman, 2005). The present study uses the non-manufacturing Z" equation continuously because the available data support its accounting

components and because changes in a common score are easier to compare across scenarios than binary classifications.

This choice has limits. Coefficients and cut-offs developed in historical samples do not transfer mechanically across periods, countries, or industries. Working capital also has a different economic meaning in banks, utilities, retailers, and software companies. Ohlson (1980) and subsequent distress research show that model specification, sampling, and outcome definitions materially affect apparent performance. Accordingly, the Z'' measure here is an accounting-resilience proxy, not an externally validated bankruptcy model.

Using changes in the continuous score partially reduces, but does not eliminate, those concerns. Each firm is compared with its own 2025 baseline under common transformations, so the analysis focuses on directional deterioration rather than literal classification. The transformation can still affect sectors differently because balance-sheet structures differ. Sensitivity tests excluding Financials and proxy-supported inputs are therefore more informative than relying on one full-sample cut-off.

2.3 Supply-Chain Shock Propagation

Production networks transmit shocks because firms depend on flows of goods, services, and revenue. Theoretical work shows that network structure can absorb local disturbances yet amplify sufficiently large or widespread shocks (Acemoglu, Carvalho, Ozdaglar, and Tahbaz-Salehi, 2012; Acemoglu, Ozdaglar, and Tahbaz-Salehi, 2015). Empirical studies find that firm-specific disruptions affect customers and suppliers when inputs are specialised, relationships are concentrated, or substitution is difficult (Barrot and Sauvagnat, 2016; Carvalho et al., 2021).

Propagation can operate in both directions. Downstream customers may lose production when a supplier fails to deliver a specialised input; upstream suppliers may lose revenue when a customer cuts purchases. Because the filings used here are major-customer disclosures made by suppliers, the observable mechanism is customer-to-supplier demand exposure. The study does not infer physical input flows or assume that a named customer is the supplier's only route to market.

Customer dependence creates a particularly relevant upstream demand channel. A supplier can lose sales even when its own facilities remain operational if a major customer reduces activity. Public filings sometimes disclose material customers and the share of revenue associated with them. Those disclosures can support auditable exposure tracing, but they are selective: firms usually report only relationships above materiality thresholds, private and foreign counterparties may not match the study universe, and a named customer does not reveal contract structure, substitutability, or the complete physical supply chain.

Customer concentration and the network measure are not independent concepts in this study. The network exposure is constructed by multiplying customer shocks by sales-share weights, so a positive association between maximum customer share and network stress is partly mathematical. Such an association describes how the measure behaves; it is not independent evidence that concentration causally increases realised climate losses.

2.4 Network Learning and Diagnostic Controls

Graph models represent firms as nodes and commercial relationships as edges. Message passing combines a firm's own attributes with information from connected firms. This is useful when the target depends on relational structure, but sparse or selectively disclosed graphs can be unrepresentative. Graph models can also appear accurate when node features already contain close transformations of the target. Reviews of graph neural networks therefore emphasise comparable feature sets, appropriate baselines, and careful validation (Wu et al., 2021).

No-edge and shuffled-edge controls are especially important. A self-only model tests whether architecture and node features explain performance without inter-firm information. A shuffled graph tests whether any topology performs as well as the observed commercial graph. If the observed graph fails to beat these controls, high R^2 shows reconstruction of the accounting target rather than meaningful network learning. This concern is consistent with broader warnings about leakage and weak evaluation in graph-learning benchmarks (Shchur et al., 2018).

Interpretability also differs across network methods. A weighted exposure sum can be inspected edge by edge, and its units remain percentage points of modeled customer stress. A multi-layer graph model can represent nonlinear interactions but may obscure whether a result comes from a firm's own features, neighbours, normalization, or architecture. In a sparse and selectively disclosed network, the simpler calculation has an evidential advantage even if a learned model produces a lower reconstruction error.

2.5 Research Gap and Added Contribution

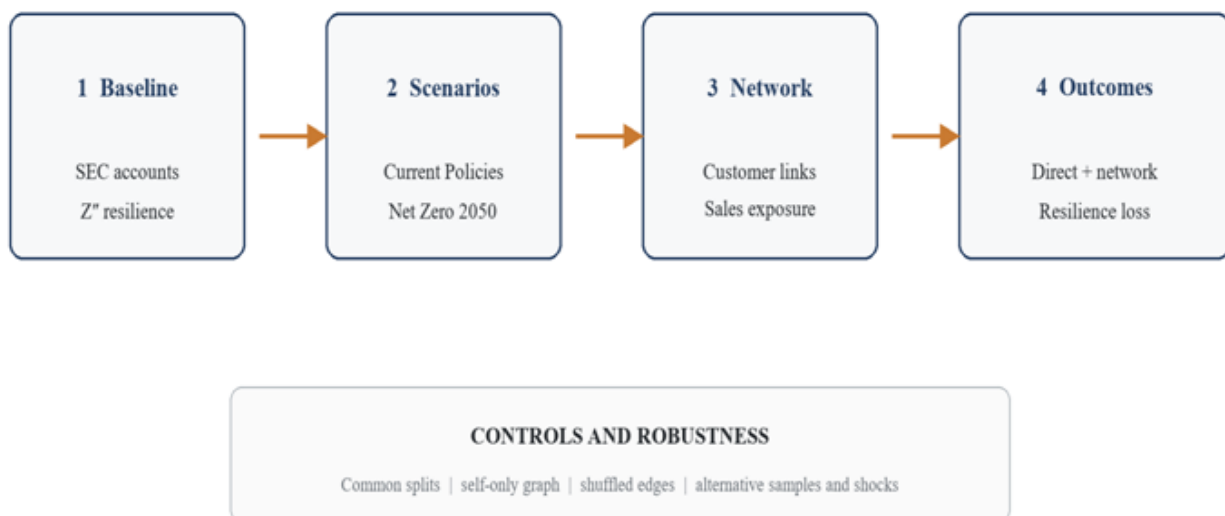
Previous studies have established four relevant points: climate exposures can affect corporate finances; scenario analysis can organise forward-looking stress tests; financial ratios can summarise current resilience; and economic shocks can propagate through production networks. Regulatory and academic work has also developed increasingly sophisticated climate and graph models. However, these streams do not automatically yield a transparent, public-data test of how disclosed material-customer dependence changes accounting-resilience estimates across the S&P 500.

This study adds three elements. First, it connects SEC financial statements to auditable 10-K customer disclosures rather than relying on a proprietary supply-chain database. Second, it separates direct sector stress from a deterministic customer-weighted exposure so that the network contribution can be traced relationship by relationship. Third, it reports strict-extraction, no-fallback, threshold, and graph-placebo tests, including results that weaken the main network claim. The purpose is not to introduce a new default predictor. It is to show what can and cannot be inferred from disclosed customer relationships under explicit scenario assumptions.

The gap is consequently methodological as well as empirical. Earlier work supplies evidence about climate pricing, scenario design, distress ratios, and production-network propagation, but those findings do not resolve how a public-data study should combine them without overstating precision. The present framework contributes a reproducible chain from source record to constructed assumption to scenario output, with explicit tests of where that chain is fragile.

3. Methodological Framework

Figure 1. Sequential research framework



Note. Source: Study design. The framework separates observed accounts, constructed scenario inputs, disclosed relational exposure, and conditional outputs.

The study is quantitative, secondary-data based, scenario based, non-experimental, and non-causal. It compares a 2025 accounting baseline with conditional scenario observations for 2030, 2040, and 2050. Figure 1 summarises the sequence: establish baseline resilience, construct sector shocks, transmit customer stress through the observed disclosed network, and compare outcomes and diagnostics.

The unit of analysis is the firm for the baseline and the firm-scenario-year for forward comparisons. The design does not estimate treatment effects because neither scenario is observed and firms are not randomly assigned to climate pathways or customer networks. Results are cross-sectional comparisons generated under common assumptions. This design is suitable because it makes the change created by adding disclosed customer information directly observable within the model.

3.1 Baseline Financial Resilience

Baseline financial resilience is measured using the non-manufacturing Altman Z'' specification:

$$Z''_i = 6.56 \frac{WC_i}{TA_i} + 3.26 \frac{RE_i}{TA_i} + 6.72 \frac{EBIT_i}{TA_i} + 1.05 \frac{BVE_i}{TL_i} \quad (1)$$

Immediately below Equation (1), WC is working capital; TA is total assets; RE is retained earnings; EBIT is earnings before interest and taxes or the documented operating-income substitute; BVE is book-value equity; TL is total liabilities; and i identifies the firm. Higher Z'' values indicate stronger accounting resilience under this construction.

Financial inputs follow a source-first hierarchy. Working capital equals reported current assets minus reported current liabilities when both are available; otherwise, assets minus liabilities is used and flagged as a proxy. Reported retained earnings and EBIT or operating income take precedence; missing retained earnings are proxied by book equity, and missing EBIT is proxied as 10 percent of revenue. Book equity and liabilities use the selected annual statement values. The score remains continuous. Conventional zones are descriptive reference points, not observed default labels.

Each component is retained separately so that an unusual score can be traced to liquidity, accumulated earnings, profitability, or leverage. Ratios with nonpositive asset or liability denominators are treated as unavailable rather than forced to zero. Provenance indicators identify whether every term was sourced or whether at least one proxy entered the formula.

3.2 Construction of Climate-Scenario Shocks

The scenario table is researcher constructed. NGFS publications provide the qualitative ordering of pathways - stronger early transition under Net Zero 2050 and greater accumulated physical risk under Current Policies - but the study does not extract official NGFS firm-level loss rates. The calibration translates that ordering into transparent sector-year inputs using three layers: a scenario path, sector sensitivities, and an explicit formula.

For scenario s and year t , the path specifies a carbon-price index, temperature change from the 2025 baseline, and an economy-wide gross-value-added adjustment. These values were selected to reproduce the relative timing and direction of the NGFS narratives, not to claim point accuracy. Each GICS sector is then mapped to Energy, Power, Industry, Agriculture, or Services. Researcher-assigned carbon intensity, transition sensitivity, and physical sensitivity rank sectors according to exposure mechanisms discussed in NGFS (2024a, 2024b), TCFD (2017), Battiston et al. (2017), and Acharya et al. (2023). Energy and Power receive high transition sensitivity because of carbon-intensive production and policy exposure. Agriculture receives the highest physical sensitivity because of dependence on temperature, water, and biological conditions. Industry receives intermediate values because of energy use, fixed capital, and supply-chain dependence. Services receive the lowest values because average direct emissions and physical asset intensity are lower, while recognising considerable within-category heterogeneity.

The carbon-intensity index is relative rather than an estimate of tonnes emitted by a specific firm. Energy is the normalization point at 150, Power is 130, Industry is 95, Agriculture is 70, and Services is 18. Transition sensitivities range from 0.45 in Energy to 0.08 in Services; physical sensitivities range from 0.42 in Agriculture to 0.08 in Services. These parameters determine relative sector ordering. They are not estimated from the S&P 500 sample and should not be used outside this experiment without recalibration.

The constructed transition loss, physical loss, and total sector shock are:

$$C_{g,s,t} = G_{s,t} - \frac{c_{s,t}}{100} \tau_g \frac{\kappa_g}{150} - \Delta Temp_{s,t}^2 \varphi_g \quad (10)$$

Immediately below Equation (2), C is the signed total sector shock; G is the selected gross-value-added adjustment; T is transition loss; P is physical loss; c is the scenario carbon-price index; τ is sector transition sensitivity; κ is the sector carbon-intensity index; $\Delta Temp$ is temperature change from the baseline; φ is sector physical sensitivity; g identifies the climate-sector group; s identifies the scenario; and t identifies the year. The constants 100, 150, and 10 scale the constructed components into percentage-point magnitudes. The firm-level direct loss is $D = \max(0, -C)$.

Table 1 reports the resulting loss magnitudes. The table is a disclosure of assumptions, not evidence that these percentages will occur. The central calibration is also multiplied by 0.75 and 1.25 in sensitivity tests. Because proportional scaling mechanically preserves many rankings, those tests assess magnitude dependence but do not validate the selected central values.

Table 1. Constructed climate-sector calibration and resulting loss paths

Climate sector	GICS mapping	κ	τ	ϕ	Current Policies 2030/40/50	Net Zero 2050 2030/40/50	Assumption rationale
Energy	Energy	150	0.45	0.30	2.69 / 6.98 / 13.94	3.16 / 2.98 / 2.64	Highest carbon intensity and transition sensitivity
Power	Utilities	130	0.40	0.35	2.85 / 7.49 / 15.09	3.06 / 2.85 / 2.51	High policy exposure and physical asset intensity
Industry	Materials; Industrials	95	0.30	0.22	2.34 / 6.00 / 11.88	2.82 / 2.41 / 1.88	Energy use, fixed capital, and production dependence
Agriculture	Consumer Staples	70	0.18	0.42	3.03 / 8.16 / 16.63	2.77 / 2.47 / 2.04	Highest sensitivity to heat, water, and biological conditions
Services	All remaining GICS sectors	18	0.08	0.08	1.79 / 4.39 / 8.43	2.54 / 1.92 / 1.18	Lower average direct emissions and asset intensity

Note. Source: Researcher construction informed by NGFS (2024a, 2024b), TCFD (2017), Battiston et al. (2017), and Acharya et al. (2023). κ = carbon-intensity index; τ = transition sensitivity; ϕ = physical sensitivity. Loss paths are percentages for 2030, 2040, and 2050. They are assumptions, not forecasts or raw NGFS outputs.

3.3 Accounting Transformation and Worked Example

For a loss rate d expressed as a proportion, the multiplier is $m = \max(0.01, 1 - d)$. Revenue, total assets, current assets, and EBIT are multiplied by m . This common proportional treatment assumes that lower activity affects sales, productive asset value, liquid operating assets, and operating earnings over the same scenario interval. It is a simplifying comparative stress rule, not an estimated behavioural equation.

Total liabilities and current liabilities are held constant. The reason is to model a short-run adverse balance-sheet state in which contractual obligations do not automatically fall when revenue or assets decline. Equity and retained earnings absorb the reduction in total assets, preserving the accounting intuition that losses first reduce owners' residual claim and accumulated earnings. This assumption is deliberately conservative for leverage, but it omits debt repayment, refinancing, new equity, covenant renegotiation, insurance recoveries, and management adaptation. Alternative liability responses would require a separate financing model not supported by the current data.

Applying the same multiplier to revenue and EBIT preserves the starting operating margin. Applying it to total and current assets assumes that the scenario loss affects the scale or value of productive and liquid assets proportionally. These rules make the transformation reproducible but exclude margin-specific cost responses and changes in asset composition. The worked example exposes the mechanics so readers can judge whether the assumptions suit a particular application.

Table 2. Worked example of a 10 percent accounting shock

Variable	Baseline	Transformation	Shocked value	Role in Z''
Revenue	1,000	× 0.90	900	Supports EBIT when proxied
Total assets	800	× 0.90	720	Denominator for three ratios
Total liabilities	400	Held constant	400	Contractual obligation
Current assets	300	× 0.90	270	Working-capital numerator
Current liabilities	200	Held constant	200	Working-capital numerator
Retained earnings	150	– 80 asset loss	70	Accumulated earnings
EBIT	80	× 0.90	72	Operating profitability
Book equity	400	– 80 asset loss	320	Residual balance-sheet claim
Calculated Z''	3.153	Baseline – shocked	2.467	Resilience loss = 0.687

Note. Source: Study illustration in arbitrary monetary units. The example explains the implemented transformation and does not represent an observed firm.

Table 2 gives a purely illustrative example in monetary units. A 10 percent shock changes revenue from 1,000 to 900, assets from 800 to 720, current assets from 300 to 270, and EBIT from 80 to 72. Liabilities remain 400 and current liabilities remain 200. The 80-unit asset reduction lowers equity from 400 to 320 and retained earnings from 150 to 70. Under Equation (1), Z'' declines from 3.153 to 2.467, so scenario-induced resilience loss is 0.687 Z-score units. This example explains the mechanics; it is not an observation for any named firm.

Scenario-induced resilience loss is defined as:

$$L_{i,s,t}^Z = Z_{i,2025}'' - \hat{Z}_{i,s,t}'' \quad (3)$$

Immediately below Equation (3), L is the scenario-induced resilience loss in Z-score units; Z''(i,2025) is firm i's observed baseline accounting score; the hatted Z''(i,s,t) is the score recalculated after applying scenario s in year t. A positive value indicates lower modeled

accounting resilience. The earlier project label "Climate VaR" is retained only for traceability and is not monetary value at risk.

3.4 Construction of the Observed Disclosed Network

Relationship construction begins with each firm's latest annual 10-K. Filing HTML is converted to normalised text, and sentence windows containing customer, sales, or revenue language and a percentage are retained. Two rule sets are applied. Strict patterns require a named customer, a revenue-action phrase, and an adjacent percentage. Broader patterns search eligible sentence windows for aliases of S&P 500 firms and attach the most prominent percentage in the window. The raw customer string, percentage, filing URL, and surrounding disclosure snippet are retained.

Customer names are normalised for case, punctuation, whitespace, legal suffixes, and selected aliases. Exact alias matches receive a name-match score of 100. Multi-word unmatched names are compared with roster aliases using token-set similarity; matches below 88 out of 100 are rejected. The 88 threshold was selected as a conservative balance between recognising minor legal-name variation and avoiding forced matches. A relationship is retained only when both supplier and customer are in the selected roster, the tickers differ, and the customer can be matched. Duplicate supplier-customer pairs are consolidated.

Consolidation counts supporting rows, takes the highest valid sales percentage, and retains the strongest confidence evidence for each directed pair. Taking the maximum is reasonable when repeated filing text contains several historical percentages or duplicated passages, but it can overstate exposure if a broad rule attaches an unrelated percentage. That risk motivates the strict-extraction sensitivity and retention of disclosure snippets for manual audit.

The extraction confidence begins at 0.92 for a strict pattern and 0.70 for a broad rule. The name-match score divided by 1,000 is added and the result is capped at 0.99. Thus an exact alias match produces confidence 0.99 for a strict extraction and 0.80 for a broad extraction. The retained graph happens to range from 0.80 to 0.99. A sensitivity cut-off of 0.80 therefore removes no principal edges and is reported as an uninformative mechanical check, not corroborating evidence.

When at least one valid customer-sales percentage between 0 and 100 is available for a supplier-customer pair, the maximum valid percentage divided by 100 becomes the edge weight. When no valid percentage remains, the fallback weight is 5 percent of the confidence score, bounded between 0.5 and 5 percent. At the observed minimum confidence of 0.80, the fallback is therefore 4 percent; at 0.99 it is 4.95 percent. The fallback is intentionally smaller than most material-customer disclosures, but it remains a researcher-created assumption.

The upstream exposure for supplier i is:

$$N_{i,s,t} = \sum_j w_{j \rightarrow i} D_{j,s,t} \quad (4)$$

Immediately below Equation (4), N is the customer-linked exposure in percentage points; $w(j \rightarrow i)$ is the disclosed or fallback sales weight from customer j to supplier i ; $D(j,s,t)$ is customer j 's direct scenario loss; i is the supplier; j indexes observed disclosed customers; s is the scenario; and t is the year. The sum is an exposure calculation, not an empirically estimated multiplier.

The graph is described throughout as the observed disclosed network. It is not the complete S&P 500 supply chain. Public filings disclose selected material customers, not every customer or supplier. The graph omits unmatched private and foreign firms, lower-tier links, undisclosed relationships, contract terms, supplier substitution, and geographic production flows. A zero network value means that no usable disclosed S&P 500 customer link was observed; it does not mean that the firm has no supply-chain risk.

3.5 Robustness and Secondary Machine-Learning Diagnostic

Robustness tests address distinct questions. Removing the 69 confidence-fallback edges tests whether results depend on researcher-assigned weights. Restricting the sample to strict extraction patterns with valid disclosed percentages tests sensitivity to the broad relationship rule. Equal 10 percent weights test whether ranking depends on a few large reported shares. Scaling shocks to 75 and 125 percent tests magnitude dependence. Fully sourced and nonfinancial samples test financial provenance and sector comparability. Hidden-vulnerability screens vary the direct-loss cut-off from 1.5 to 2.5 percent and the network-tail threshold from the 85th to the 95th percentile.

The GNN is secondary. A residual graph model, a financial-only comparator with the same non-network features, a self-only graph, and shuffled-edge placebos use common train, validation, and test node splits across five seeds. RMSE and R^2 measure reconstruction of the accounting target only. Because the target is calculated from variables closely related to the model inputs, high R^2 can arise without learning economically meaningful network effects. Full architecture, hyperparameters, and results are moved to the supplementary material.

Mathematically expected and evidentially informative checks are distinguished. Proportional shock scaling is expected to change magnitudes proportionally, so a high correlation after scaling confirms implementation consistency but supplies little new economic evidence. Removing fallback edges or restricting extraction rules changes the information set and directly tests whether substantive conclusions depend on researcher-created relationship assumptions.

4. Data

The analytical population is the selected current S&P 500 roster. It contains 503 securities because the index can include more than one listed share class for a constituent company. The baseline year is 2025. Current Policies and Net Zero 2050 are evaluated in 2030, 2040, and 2050, producing six scenario-year observations per firm and 3,018 firm-scenario-year rows.

Financial data come primarily from SEC EDGAR Company Facts XBRL and the latest annual 10-K. Revenue, assets, liabilities, equity, current assets, current liabilities, retained earnings, and EBIT or operating income support the accounting-resilience calculation. A documented yfinance value is used only when a comparable SEC item cannot be constructed. Major-customer disclosures and their filing snippets supply the relationship evidence. NGFS reports provide scenario narratives, while all numerical sector shocks are constructed by the researcher using the method in Section 3.2.

The latest usable annual fact is selected from comparable annual periods associated with the relevant filing. Where more than one standardized XBRL concept can represent an item, the pipeline follows a documented precedence order and retains the selected source. Financial values are not interpolated across scenario years; future observations are created only by the stated scenario transformation.

Table 3. Data sources and analytical role

Data component	Source	Coverage	Principal use	Qualification
Firm roster and sectors	Selected S&P 500 roster	503 securities; 2025 baseline	Population and GICS mapping	Includes multiple listed share classes
Financial statements	SEC EDGAR Company Facts XBRL and 10-K	Latest usable annual facts	Accounting-resilience inputs	Preferred source
Fallback financial values	yfinance	Unavailable comparable fields only	Reduce avoidable missingness	Documented; not preferred
Customer relationships	Major-customer 10-K disclosures	Latest annual filing	Directed customer-to-supplier edges	Observable material-customer network only
Climate scenarios	NGFS narratives; study calibration	2030, 2040, 2050	Transition and physical stress inputs	NGFS-inspired; not raw NGFS extraction

Note. Sources: SEC (2025a, 2025b), NGFS (2024a, 2024b), and study construction.

Company names, tickers, Central Index Keys, and sectors are standardised before merging. Missing values are not automatically coded as zero. Field-level source labels identify SEC values, documented fallbacks, and accounting proxies. Relationship cleaning normalises customer names, applies the matching rules described in Section 3.4, removes self-links, and consolidates duplicate supplier-customer pairs.

Coverage is strongest for revenue, assets, equity, and retained earnings. Revenue is SEC-sourced for 498 firms, while assets and equity are SEC-sourced for 502. Retained earnings are SEC-sourced for 494 firms and EBIT or operating income for 416. Liabilities are SEC-sourced for 372 firms, with the remainder relying on documented fallback values. Current assets and current liabilities are available from the preferred source for 423 firms. These differences explain why 460 firms, rather than all 503, are classified as having fully sourced *Z*" inputs.

Table 4. Final sample and network coverage

Measure	Value	Interpretation
S&P 500 securities	503	Analytical roster
Baseline year	2025	Observed financial baseline
Scenario years	2030, 2040, 2050	Two scenarios in each year
Extracted relationship rows	1,684	Before matching and deduplication
Usable directed edges	612	Customer-to-supplier relationships
Firm-scenario-year observations	3,018	503 x 2 scenarios x 3 years

Note. Source: Study analytical database. Counts are reconciled before model estimation.

The relationship pipeline starts with 1,684 extracted rows and produces 612 unique usable edges connecting 375 suppliers to 75 S&P 500 customers. Of those edges, 543 have a valid percentage-derived weight and 69 use the confidence fallback. The strict-extraction robustness sample is much smaller: 27 edges connect 25 suppliers to 11 customers. This difference is important because it demonstrates that the broad extraction rule expands coverage at the cost of greater measurement risk.

Four evidential categories are maintained. SEC accounts and filing excerpts are directly observed source data. yfinance substitutions, accounting proxies, and confidence-fallback edge weights are documented supported values. GICS mappings, sensitivity parameters, scenario values, and screen thresholds are researcher-constructed inputs. Direct losses, network exposure, *Z*" changes,

rankings, and model diagnostics are derived results. The final data are suitable for transparent comparative stress analysis but not for reconstructing the complete production network or forecasting realised firm outcomes.

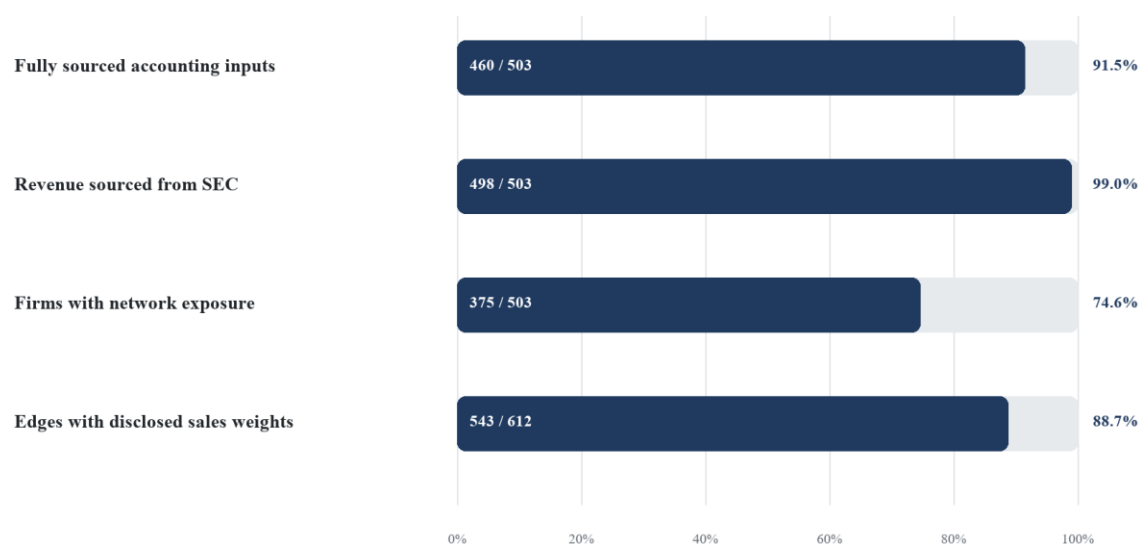
The current roster introduces membership selection because it represents firms included at the collection date rather than historical or future index membership. Multiple share classes remain separate securities where the roster supplies distinct identifiers. These choices preserve reproducibility but limit generalisation to firms that entered, exited, merged, or failed outside the selected roster.

5. Results and Discussion

5.1 Data Coverage and Baseline Resilience

The final database contains 503 firms and 3,018 firm-scenario-year observations. Revenue is sourced from SEC records for 498 firms and from a documented fallback for five. Assets and equity are SEC-sourced for 502 firms. Across the complete Z'' formula, 460 firms have fully sourced inputs and 43 require at least one supported proxy or fallback construction. Figure 2 distinguishes accounting coverage from network reach.

Figure 2. Data coverage and source quality



Note. Source: Study provenance and data-quality records. Denominators are firms for the first three measures and usable edges for the fourth.

Economically, the coverage result means that the common resilience measure is available across nearly the full roster, but the indirect-exposure analysis applies only where public filings reveal a usable relationship. An unconnected firm is not necessarily independent; it is simply unobserved in this disclosed network.

The customer side is also narrow: 75 matched S&P 500 customer nodes appear in the principal graph. Large, widely named customers can recur across filings, while suppliers dependent on private, foreign, or generically described customers receive no comparable edge. Network exposure may therefore reflect disclosure practice and roster overlap as well as economic dependence.

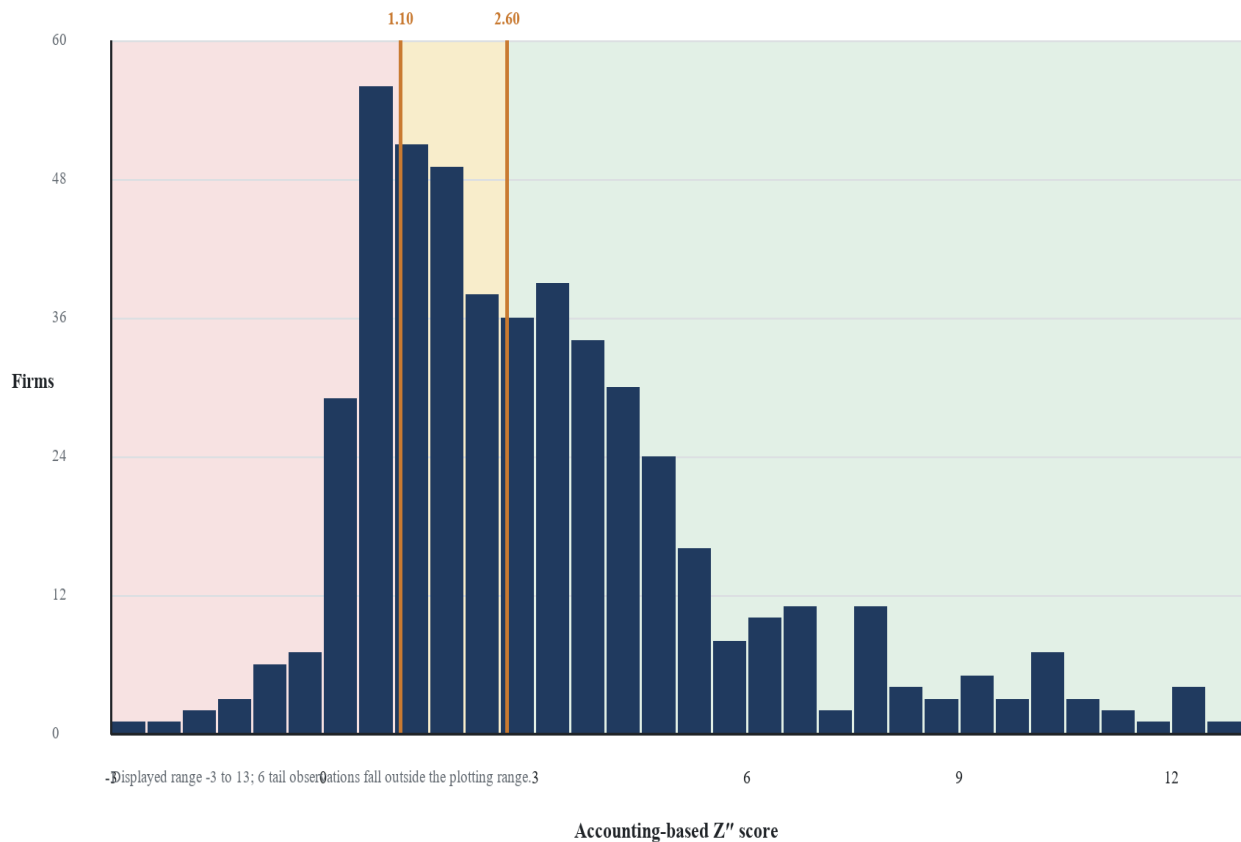
The corrected baseline Z'' score has a mean of 3.13, a median of 2.63, and a standard deviation of 3.23. Descriptive zones place 120 firms below 1.10, 129 between 1.10 and 2.60, and 254 above 2.60. Table 5 and Figure 3 summarise the distribution.

Table 5. Descriptive statistics

Measure	N	Mean / share	Median	SD
Corrected Z''	503	3.132	2.631	3.225
Descriptive zone: below 1.10	503	23.9%	-	-
Descriptive zone: 1.10-2.60	503	25.6%	-	-
Descriptive zone: above 2.60	503	50.5%	-	-
Fully sourced formula inputs	503	91.5%	-	-
Connected supplier firms	503	74.6%	-	-

Note. Source: Study calculations. Zones are descriptive and are not observed defaults. SD = standard deviation.

Figure 3. Distribution of baseline accounting-based resilience



Note. Source: Study calculations; $N = 503$. Reference zones are descriptive, not observed bankruptcy outcomes.

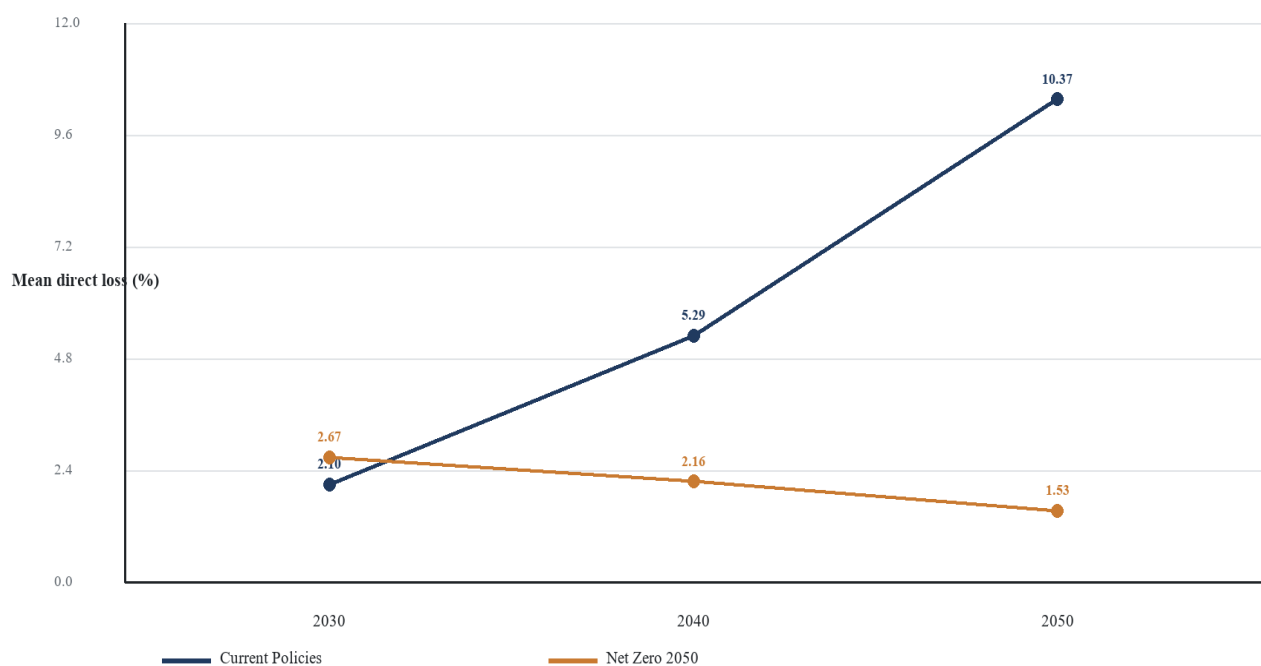
These categories describe current accounting ratios, not future bankruptcies. Results are similar among the 460 fully sourced firms and after excluding Financials, which reduces concern that one proxy drives the broad baseline pattern. Cross-sector comparisons nevertheless remain descriptive because the same accounting ratio can have different meanings across business models.

Utilities and Real Estate have comparatively low sector medians, while Industrials, Health Care, and Information Technology have higher medians. The differences may reflect capital intensity, liability structure, and accumulated earnings rather than an underlying probability of failure. The baseline is most useful as the starting condition against which the same firm's modeled score changes.

5.2 Direct Scenario Effects

Under Current Policies, mean direct loss rises from 2.10 percent in 2030 to 5.29 percent in 2040 and 10.37 percent in 2050. Under Net Zero 2050, mean direct loss is 2.67 percent in 2030, 2.16 percent in 2040, and 1.53 percent in 2050. Figure 4 shows the reversal.

Figure 4. Direct climate stress by scenario and year



Note. Source: Study-authored NGFS-inspired calibration; $N = 503$ firms per scenario-year. Values are mean direct percentage losses.

The economic interpretation is the timing trade-off built into the calibration. Net Zero 2050 imposes more near-term adjustment, whereas Current Policies permits physical stress to accumulate. The result is mathematically implied by the selected scenario path and should be read as a controlled comparison, not empirical evidence that either path will deliver the reported percentages.

Sector differences also follow the constructed sensitivities. Energy and Power remain relatively exposed to transition pressure, Agriculture becomes highly exposed to accumulated physical stress, Industry is intermediate, and Services is lower on average. These are category-level

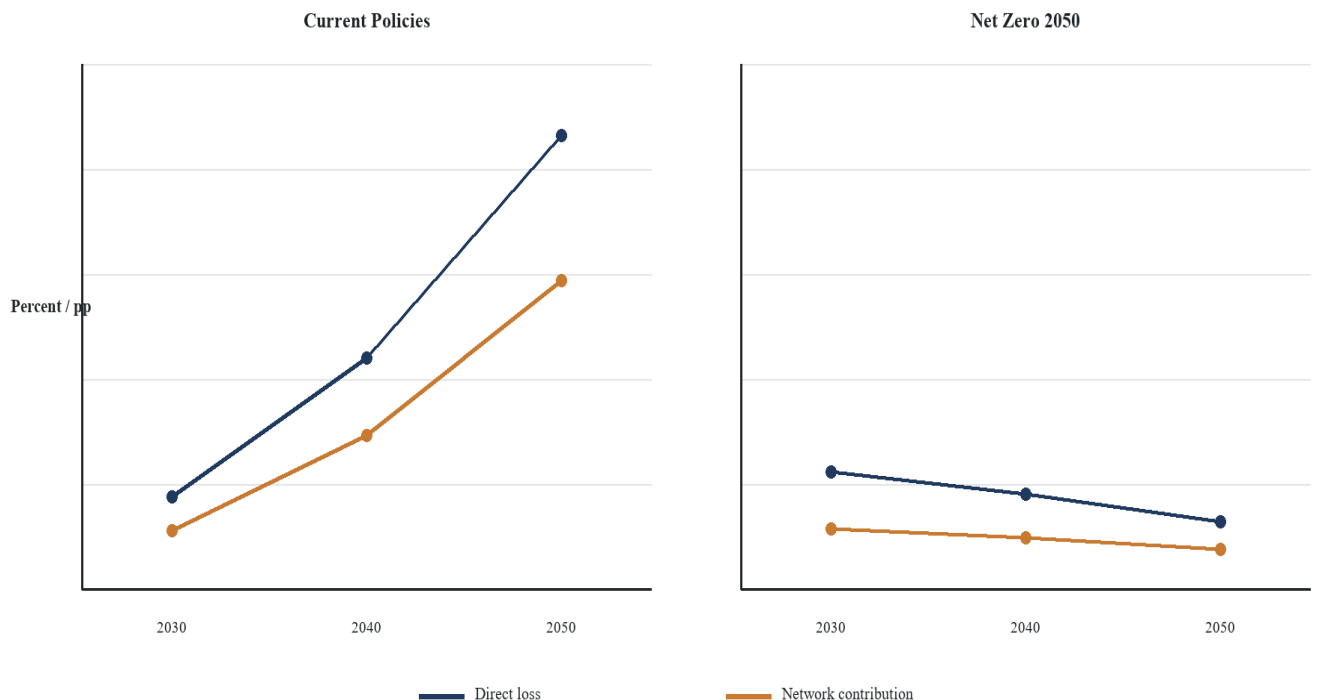
assumptions; individual firms can differ substantially in emissions, location, technology, insurance, and adaptation.

The scenario-year comparison is therefore more defensible than a claim about exact cross-sector loss. It shows how one explicit calibration changes through time and provides a basis for checking whether network connections alter the ordering. It does not establish the welfare effect of either pathway or include benefits from avoided damages outside corporate accounts.

5.3 Network Exposure and Customer Concentration

In the principal 612-edge graph, mean customer-linked exposure under Current Policies rises from 1.33 percentage points in 2030 to 3.50 in 2040 and 7.04 in 2050. Under Net Zero 2050, it declines from 1.38 to 0.90 percentage points. Figure 5 and Table 6 report the direct and customer-linked components separately.

Figure 5. Direct and network-amplified stress



Note. Source: Study calculations; $N = 503$, of which 375 have disclosed network exposure. Network values are additional percentage points.

Table 6. Scenario and network results

Scenario	Year	Direct loss (%)	Network contribution (pp)	Network share	Connected firms
Current Policies	2030	2.10	1.33	38.8%	375
Current Policies	2040	5.29	3.50	39.8%	375
Current Policies	2050	10.37	7.04	40.5%	375
Net Zero 2050	2030	2.67	1.38	34.1%	375
Net Zero 2050	2040	2.16	1.17	35.2%	375
Net Zero 2050	2050	1.53	0.90	37.1%	375

Note. Source: Study calculations; $N = 503$ firms per scenario-year. pp = percentage points. Network share equals network contribution divided by direct plus network exposure.

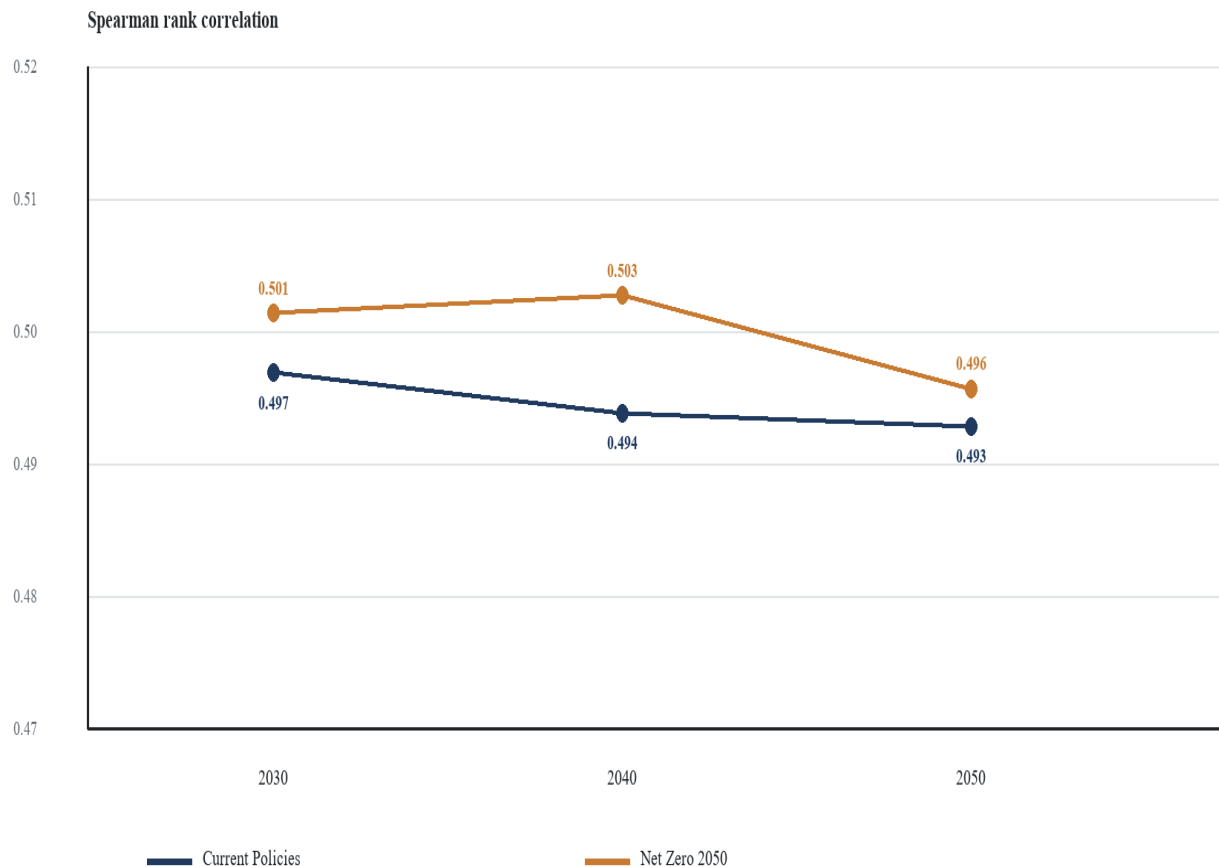
For a connected supplier, the network component answers a practical question: how much additional scenario exposure is attached to the disclosed customers on which revenue depends? It is not an observed sales decline. Its size is mechanically determined by customer shocks and edge weights, and the static graph does not allow relationships to change.

The network affects the same connected firms in every scenario-year because topology is fixed. Only customer shock intensity changes. This makes pathways comparable but excludes supplier switching, customer exit, new contracts, vertical integration, and adaptation between 2025 and 2050. The reported network share is a decomposition of the model, not a historical multiplier.

Maximum disclosed customer share and network exposure have Spearman correlations of approximately 0.49 to 0.50 across scenario-years. Figure 6 visualises the relationship.

This correlation should not be treated as independent validation. Customer sales shares are used directly to construct network exposure, so part of the association is built into the measure. The result confirms that concentration affects the screen as designed; it does not establish a causal effect on realised losses.

Figure 6. Customer concentration and network loss



Note. Source: Study calculations; $N = 375$ connected firms per scenario-year. Values are Spearman correlations with maximum disclosed customer sales exposure.

5.4 Relationship-Quality and Other Robustness Tests

Removing all 69 confidence-fallback edges leaves 543 edges and 353 connected suppliers. Average Current Policies network exposure in 2050 changes only from 7.044 to 6.970 percentage points, and rank correlations with the principal specification remain approximately 0.9999. This test directly addresses fallback weights and indicates that those 69 edges are not responsible for the main ordering.

The stricter extraction test gives a much more cautious result. Requiring a strict customer-revenue pattern and a valid disclosed percentage leaves only 27 edges, connecting 25 suppliers to

11 customers. Mean Current Policies network exposure in 2050 falls to 0.125 percentage points across all 503 firms, and mean Net Zero 2050 exposure falls to 0.017 percentage points. Rank correlations with the main vulnerability score remain about 0.95 to 0.96 because direct and financial components still dominate the composite ranking. The large reduction in network magnitude shows that the principal indirect-exposure result is sensitive to the broader relationship-extraction rule.

This is the most consequential robustness finding. The no-fallback test shows that the numeric fallback formula is not driving the result, but the strict-only test shows that relationship identification is. Uncertainty arises mainly from deciding whether a sentence establishes a customer-revenue link, not from the small weight assigned when a valid percentage is absent. Future applications should prioritise manual validation or structured disclosure rather than further tuning of the fallback formula.

Equal 10 percent edge weights produce rank correlations of roughly 0.965 to 0.967 with the main score. Scaling climate shocks to 75 or 125 percent changes losses proportionally and leaves rankings unchanged. These findings are largely expected from the model's linear and within-scenario normalisation. They demonstrate computational stability, not external validation of the shock level.

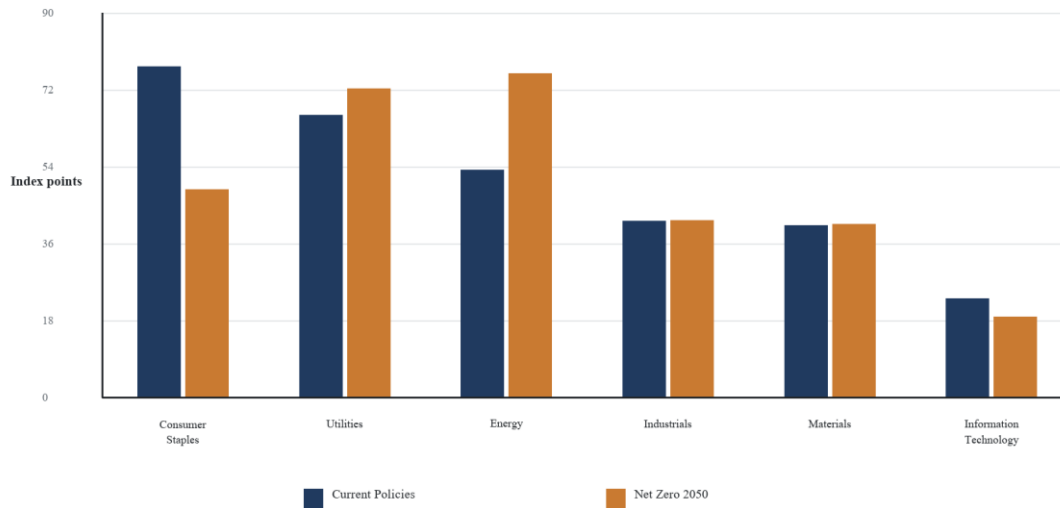
The hidden-vulnerability screen is also threshold dependent. For Net Zero 2050 in 2050, the principal rule - direct loss below 2 percent and network exposure at or above the connected-firm 90th percentile - identifies 33 firms. Across direct-loss cut-offs of 1.5, 2.0, and 2.5 percent and network thresholds at the 85th, 90th, and 95th percentiles, counts range from 11 to 55. The screen is therefore useful for prioritising cases but should not be interpreted as a stable estimate of population prevalence.

Fully sourced and nonfinancial samples produce baseline means and descriptive-zone shares similar to the principal sample. This supports broad comparative use of the accounting measure, while the narrow strict relationship sample prevents an equally strong statement about the magnitude of indirect exposure.

5.5 Sector Rankings and Hidden Vulnerability

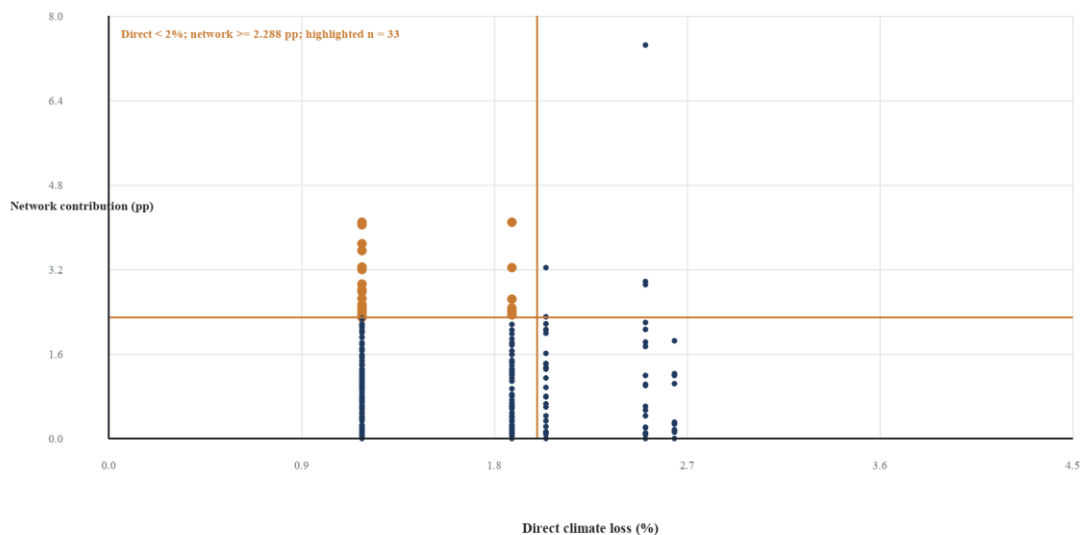
Network adjustment can reorder firms that share the same direct sector shock because their disclosed customers differ. Figure 7 compares selected sector vulnerability scores in 2050, and Figure 8 shows the hidden-vulnerability screen.

Figure 7. Selected sector vulnerability in 2050



Note. Source: Study deterministic vulnerability index. Values are relative index points, not percentages or default probabilities.

Figure 8. Hidden-vulnerability screen under Net Zero 2050



Note. Source: Study calculations; $N = 503$. Highlighted firms have direct loss below 2 percent and network exposure at or above the connected-firm 90th percentile.

The economic value of the screen lies in disagreement between two views: a firm's own sector appears lightly stressed, but its disclosed customers are highly stressed. Such cases warrant review of contracts, substitution options, customer diversification, insurance, and geographic exposure. They are not investment labels, and the 33 highlighted observations are conditional on the principal extraction and threshold rules.

Named-company rankings remain in the reproducibility files rather than the main discussion. A highly ranked company may have insurance, geographic diversification, pricing power, or alternative customers that the model cannot observe. Conversely, an unconnected company may have important dependencies below disclosure thresholds. Rankings are prompts for investigation, not conclusions about management quality or investment value.

5.6 Secondary GNN Diagnostic and Interpretation Boundary

The learned-model diagnostic does not support an incremental topology claim. Across common splits, the observed-graph model reconstructs the accounting target with mean test RMSE 0.119, but the self-only graph performs better at 0.076 and the shuffled-edge placebo is also slightly better at 0.112. High R^2 values are expected because the target is an accounting formula and closely related variables appear among the inputs.

This result means the GNN may be learning accounting relationships and architecture-specific transformations rather than meaningful supply-chain effects. It is therefore not used to support the principal conclusion. Full model details and comparison metrics appear in Supplementary Table S1. The defensible evidence is the transparent customer-weighted scenario calculation, qualified by its strong sensitivity to strict relationship extraction.

Observed accounts, constructed assumptions, and modeled outcomes remain separate. No reported scenario loss is an actual future financial outcome. No observed defaults, realised climate losses, bond spreads, or returns are used. The results support conditional scenario comparison and case screening only.

Taken together, the results answer the research question in two layers. Disclosed customer information can materially change the principal screen for connected firms, but the estimated magnitude is not robust to a strict relationship definition. The method is useful for discovering candidate dependencies and testing assumptions; stronger claims about propagation require verified relationships and observed outcomes.

6. Policy and Practical Implications

Corporate risk teams should examine material-customer dependence alongside own-sector exposure. A practical screen can apply the same scenario to the firm and its disclosed customers, then review cases where a modest direct shock coexists with concentrated customer stress. Because the strict-extraction result is much smaller, any flagged case should be verified against the original filing before decisions are made.

Risk owners should preserve the decomposition between direct sector loss, customer-linked exposure, and accounting resilience. Combining them into one opaque rating makes it difficult to identify whether a high score comes from a weak balance sheet, an aggressive scenario assumption, or a questionable relationship match. A traceable case file should record the filing sentence, customer identity, reported percentage, scenario year, and qualitative evidence about substitutability.

Investors and lenders can use the framework as a due-diligence prompt rather than a rating or price target. Relevant follow-up questions include whether the disclosed customer is still material, whether contracts protect volumes or payment terms, whether the supplier can substitute customers, and whether the customer has credible adaptation plans. The scenario output should be treated as a range of conditional exposures, not a forecast.

For portfolio analysis, the appropriate use is comparative: apply the same transparent assumptions across holdings, identify cases sensitive to network inclusion, and allocate research effort. It would be inappropriate to translate the percentage-point measure directly into expected return, credit spread, or capital requirements without a validated empirical mapping.

Standard setters could improve the evidence base through machine-readable identifiers for material customers and suppliers, consistent percentage fields, and clearer reporting of changes in commercial concentration. Better disclosure would reduce false matches and allow researchers to distinguish customer dependence from the physical flow of inputs. Supervisory exercises that incorporate network effects should also disclose how missing counterparties and incomplete tiers are handled.

The strict-extraction result gives this recommendation practical importance. When customer names and revenue shares appear only in narrative text, automated methods face a trade-off between coverage and precision. Structured identifiers and tagged percentages would permit broader coverage without relying on sentence-level heuristics and would make year-to-year changes auditable.

Future research should replace the static calibration with validated direct ingestion of scenario variables, add geographic assets and suppliers, and test against observed outcomes such as earnings revisions, bond spreads, delivery delays, or realised disaster losses. Commercial multi-tier data could evaluate whether the public-disclosure network understates or misrepresents connectivity. More complex graph models should be considered only after stronger relational data and external labels are available.

Dynamic analysis could allow liabilities, margins, and customer relationships to respond over time. Alternative resilience measures would test dependence on Z . These extensions would address the assumptions that currently dominate uncertainty rather than adding complexity for its own sake.

7. Conclusion

Incorporating disclosed customer dependence changes the scenario-based vulnerability screen for connected firms, but the size of that change depends strongly on relationship extraction. The principal 612-edge graph produces a material customer-linked component, while a strict exact-disclosure test retains only 27 edges and substantially reduces average network exposure. Public customer data can therefore reveal potentially hidden dependencies, but coverage and parsing quality determine how much confidence can be placed in the resulting magnitude.

The direct scenario comparison is clearer: under the stated calibration, Current Policies becomes substantially more stressful by 2050, whereas Net Zero 2050 front-loads adjustment and lowers later physical stress. The GNN adds no validated evidence because observed topology does not outperform self-only or shuffled controls. All findings are conditional scenario estimates, not predictions of actual losses, default, or bankruptcy.

The most defensible conclusion is that customer relationships deserve a place in climate due diligence, but public narrative disclosures are better suited to auditable case screening than precise system-wide loss estimation.

References

- Acemoglu, D., Carvalho, V. M., Ozdaglar, A., and Tahbaz-Salehi, A. (2012). The network origins of aggregate fluctuations. *Econometrica*, 80(5), 1977–2016. <https://doi.org/10.3982/ECTA9623>
- Acemoglu, D., Ozdaglar, A., and Tahbaz-Salehi, A. (2015). Systemic risk and stability in financial networks. *American Economic Review*, 105(2), 564–608. <https://doi.org/10.1257/aer.20130456>

- Acharya, V. V., Berner, R., Engle, R., Jung, H., Stroebel, J., Zeng, X., and Zhao, Y. (2023). Climate stress testing. *Annual Review of Financial Economics*, 15, 291–326. <https://doi.org/10.1146/annurev-financial-110821-022532>
- Alogoskoufis, S., Dunz, N., Emambakhsh, T., Hennig, T., Kaijser, M., Kouratzoglou, C., Muñoz, M. A., Parisi, L., and Salleo, C. (2021). ECB economy-wide climate stress test. ECB Occasional Paper No. 281. European Central Bank. <https://www.ecb.europa.eu/pub/pdf/scpops/ecb.op281~05a7735b1c.en.pdf>
- Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance*, 23(4), 589–609. <https://doi.org/10.1111/j.1540-6261.1968.tb00843.x>
- Altman, E. I. (2005). An emerging market credit scoring system for corporate bonds. *Emerging Markets Review*, 6(4), 311–323. <https://doi.org/10.1016/j.ememar.2005.09.007>
- Aroussi, R. (2026). yfinance: Download market data from Yahoo! Finance's API. Software documentation. <https://github.com/ranaroussi/yfinance>
- Aswani, J., Raghunandan, A., and Rajgopal, S. (2024). Are carbon emissions associated with stock returns? *Review of Finance*, 28(1), 75–106. <https://doi.org/10.1093/rof/rfad013>
- Bank of England. (2022). Results of the 2021 Climate Biennial Exploratory Scenario. <https://www.bankofengland.co.uk/stress-testing/2022/results-of-the-2021-climate-biennial-exploratory-scenario>
- Barrot, J.-N., and Sauvagnat, J. (2016). Input specificity and the propagation of idiosyncratic shocks in production networks. *Quarterly Journal of Economics*, 131(3), 1543–1592. <https://doi.org/10.1093/qje/qjw018>
- Battiston, S., Mandel, A., Monasterolo, I., Schütze, F., and Visentin, G. (2017). A climate stress-test of the financial system. *Nature Climate Change*, 7, 283–288. <https://doi.org/10.1038/nclimate3255>
- Bolton, P., and Kacperczyk, M. (2021). Do investors care about carbon risk? *Journal of Financial Economics*, 142(2), 517–549. <https://doi.org/10.1016/j.jfineco.2021.05.008>
- Bronstein, M. M., Bruna, J., Cohen, T., and Veličković, P. (2021). Geometric deep learning: Grids, groups, graphs, geodesics, and gauges. *arXiv*. <https://doi.org/10.48550/arXiv.2104.13478>

- Carvalho, V. M. (2014). From micro to macro via production networks. *Journal of Economic Perspectives*, 28(4), 23–48. <https://doi.org/10.1257/jep.28.4.23>
- Carvalho, V. M., Nirei, M., Saito, Y. U., and Tahbaz-Salehi, A. (2021). Supply chain disruptions: Evidence from the Great East Japan earthquake. *Quarterly Journal of Economics*, 136(2), 1255–1321. <https://doi.org/10.1093/qje/qjaa044>
- Dietz, S., Bowen, A., Dixon, C., and Gradwell, P. (2016). Climate value at risk of global financial assets. *Nature Climate Change*, 6, 676–679. <https://doi.org/10.1038/nclimate2972>
- Financial Stability Board. (2020). The implications of climate change for financial stability. <https://www.fsb.org/2020/11/the-implications-of-climate-change-for-financial-stability/>
- Giglio, S., Kelly, B., and Stroebe, J. (2021). Climate finance. *Annual Review of Financial Economics*, 13, 15–36. <https://doi.org/10.1146/annurev-financial-102620-103311>
- Hamilton, W. L., Ying, R., and Leskovec, J. (2017). Inductive representation learning on large graphs. *Advances in Neural Information Processing Systems*, 30. <https://proceedings.neurips.cc/paper/2017/hash/5dd9db5e033da9c6fb5ba83c7a7e9-Abstract.html>
- Harris, C. R., Millman, K. J., van der Walt, S. J., et al. (2020). Array programming with NumPy. *Nature*, 585, 357–362. <https://doi.org/10.1038/s41586-020-2649-2>
- Hong, H., Karolyi, G. A., and Scheinkman, J. A. (2020). Climate finance. *Review of Financial Studies*, 33(3), 1011–1023. <https://doi.org/10.1093/rfs/hhz146>
- Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science and Engineering*, 9(3), 90–95. <https://doi.org/10.1109/MCSE.2007.55>
- Inoue, H., and Todo, Y. (2019a). Firm-level propagation of shocks through supply-chain networks. *Nature Sustainability*, 2, 841–847. <https://doi.org/10.1038/s41893-019-0351-x>
- Inoue, H., and Todo, Y. (2019b). Propagation of negative shocks across nation-wide firm networks. *PLOS ONE*, 14(3), e0213648. <https://doi.org/10.1371/journal.pone.0213648>
- Inoue, H., and Todo, Y. (2020). The propagation of the economic impact through supply chains: The case of a mega-city lockdown against the spread of COVID-19. *PLOS ONE*, 15(9), e0239251. <https://doi.org/10.1371/journal.pone.0239251>

- Intergovernmental Panel on Climate Change. (2022). Climate Change 2022: Impacts, Adaptation and Vulnerability. Cambridge University Press. <https://doi.org/10.1017/9781009325844>
- Kinnan, C., Wang, S.-Y., Wang, Y., and Zhang, Y. (2024). The economic impacts of climate risk: Evidence from Chinese firms. *American Economic Review*, 114. <https://www.aeaweb.org/articles?id=10.1257/aer.20220892>
- Kipf, T. N., and Welling, M. (2017). Semi-supervised classification with graph convolutional networks. *International Conference on Learning Representations*. <https://arxiv.org/abs/1609.02907>
- Li, Q., Shan, H., Tang, Y., and Yao, V. (2024). Corporate climate risk: Measurements and responses. *Review of Financial Studies*, 37(6), 1778–1830. <https://doi.org/10.1093/rfs/hhad094>
- McKinney, W. (2010). Data structures for statistical computing in Python. *Proceedings of the 9th Python in Science Conference*, 56–61. <https://doi.org/10.25080/Majora-92bf1922-00a>
- Merton, R. C. (1974). On the pricing of corporate debt: The risk structure of interest rates. *Journal of Finance*, 29(2), 449–470. <https://doi.org/10.2307/2978814>
- Network for Greening the Financial System. (2024a). NGFS long-term climate scenarios for central banks and supervisors. Phase V. <https://www.ngfs.net/ngfs-scenarios-portal/>
- Network for Greening the Financial System. (2024b). NGFS scenarios technical documentation. Phase V. <https://www.ngfs.net/ngfs-scenarios-portal/>
- Newman, M. E. J. (2018). *Networks* (2nd ed.). Oxford University Press. <https://doi.org/10.1093/oso/9780198805090.001.0001>
- Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, 18(1), 109–131. <https://doi.org/10.2307/2490395>
- Paszke, A., Gross, S., Massa, F., et al. (2019). PyTorch: An imperative style, high-performance deep learning library. *Advances in Neural Information Processing Systems*, 32. <https://proceedings.neurips.cc/paper/2019/hash/bdbca288fee7f92f2bfa9f7012727740-Abstract.html>
- Pedregosa, F., Varoquaux, G., Gramfort, A., et al. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830. <https://jmlr.org/papers/v12/pedregosa11a.html>

- Ranger, N., Mahul, O., and Monasterolo, I. (2022). Assessing financial risks from physical climate shocks: A framework for scenario generation. World Bank Policy Research Working Paper 9995. <https://documents.worldbank.org/en/publication/documents-reports/documentdetail/470581648469518534/>
- Schuermann, T. (2014). Stress testing banks. *International Journal of Forecasting*, 30(3), 717–728. <https://doi.org/10.1016/j.ijforecast.2013.10.003>
- Shchur, O., Mumme, M., Bojchevski, A., and Günnemann, S. (2018). Pitfalls of graph neural network evaluation. *Relational Representation Learning Workshop, NeurIPS*. <https://arxiv.org/abs/1811.05868>
- Stroebel, J., and Wurgler, J. (2021). What do you think about climate finance? *Journal of Financial Economics*, 142(2), 487–498. <https://doi.org/10.1016/j.jfineco.2021.08.004>
- Task Force on Climate-related Financial Disclosures. (2017). Recommendations of the Task Force on Climate-related Financial Disclosures. Financial Stability Board. <https://www.fsb-tcfd.org/recommendations/>
- Tokui, J., Kawasaki, K., and Miyagawa, T. (2017). The economic impact of supply chain disruptions from the Great East-Japan earthquake. *Japan and the World Economy*, 41, 59–70. <https://doi.org/10.1016/j.japwor.2016.12.005>
- U.S. Securities and Exchange Commission. (2025a). EDGAR application programming interfaces. <https://www.sec.gov/search-filings/edgar-application-programming-interfaces>
- U.S. Securities and Exchange Commission. (2025b). Inline XBRL. <https://www.sec.gov/structureddata/osd-inline-xbrl.html>
- Veličković, P., Cucurull, G., Casanova, A., Romero, A., Liò, P., and Bengio, Y. (2018). Graph attention networks. *International Conference on Learning Representations*. <https://arxiv.org/abs/1710.10903>
- World Economic Forum. (2024). The Global Risks Report 2024 (19th ed.). <https://www.weforum.org/publications/global-risks-report-2024/>
- Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., and Yu, P. S. (2021). A comprehensive survey on graph neural networks. *IEEE Transactions on Neural Networks and Learning Systems*, 32(1), 4–24. <https://doi.org/10.1109/TNNLS.2020.2978386>
- Xu, K., Hu, W., Leskovec, J., and Jegelka, S. (2019). How powerful are graph neural networks? *International Conference on Learning Representations*. <https://arxiv.org/abs/1810.00826>

Zhou, J., Cui, G., Hu, S., et al. (2020). Graph neural networks: A review of methods and applications. *AI Open*, 1, 57–81. <https://doi.org/10.1016/j.aiopen.2021.01.001>