

Unknown Is Not Zero: Heat Stress, Data-Center Power-Demand Exposure, and Utility Volatility

Fuli Yang

GARCH Institute, Harvard Square, Cambridge, MA 02138, USA

SSRN version 1.2 — September 2026

Abstract

This paper asks whether service-territory heat stress and disclosed data-center power-demand exposure jointly amplify the volatility of U.S. listed electric utilities. Its principal finding is methodological but empirically consequential: pre-disclosure silence is not a verified zero. The study constructs an auditable point-in-time chain from dated operating or binding customer-load megawatts to the regulated utility, the applicable EIA Form 861 peak-load denominator, a representative NOAA station, daily parent-company returns, and lag-available financial controls. Ten issuers satisfy a broad data-center four-quarter evidence rule over 2021–2025, although the measure is not uniformly AI-specific and one issuer is predominantly exposed to cryptocurrency hosting. Reclassifying 8,685 legacy zero-coded issuer-days as unknown leaves 3,855 positive-disclosure issuer-days and 212 complete portfolio observations across nine variance-recursion segments. The measurement path is non-monotone: the uncentered interaction is 0.143 in the legacy zero-coded panel and 0.060 when OTTR is excluded; after the common-positive restriction it reaches -2.000 at the coefficient bound, while centering yields an interior estimate of -0.118 . Before centering, $\text{Corr}(C, C \times A) = 0.979$ and the interaction VIF is 29.4; centering lowers the VIF to 1.09. In the final Student-t GARCH-X specification, AIC and BIC select the no-interaction model and a 499-replication conditional parametric bootstrap does not reject it ($LR = 1.638$, $p = 0.324$). The current design therefore fails to detect heat-by-load amplification; it does not establish that the population effect is zero. H3 remains untested, and the pre-specified Markov-switching extension is not estimated. Positively, the audit identifies how exposure coding and coordinate construction can manufacture apparent interaction evidence, supplies a reusable regulatory-data architecture, and defines an advance rule under which a longer panel can retest the same hypotheses without redefining the original evidence.

Keywords: heat stress; data centers; artificial intelligence infrastructure; electricity demand; conditional volatility; GARCH-X; measurement audit

JEL classifications: C32; C58; G12; G14; Q54

1. Introduction

Climate risk and artificial intelligence are often treated as separate sources of financial change. Their physical infrastructure intersects in the electricity system: data centers require grid connections and substantial power capacity, while extreme heat raises cooling demand and tightens operating conditions. For listed utilities, the empirical question is whether disclosed incremental data-center load changes how local heat stress is translated into return volatility.

The scale of the physical connection has changed quickly. Berkeley Lab estimates that U.S. data centers consumed about 176 TWh in 2023, approximately 4.4% of national electricity consumption, and projects a wide 2028 range of 325–580 TWh, or 6.7–12.0% of U.S. electricity use (Shehabi et al., 2024). The width of that range is itself economically relevant. Utilities must plan generation, transmission, distribution, and financing before realized demand is known, while duplicated interconnection requests and uncertain completion rates can inflate pipeline measures. For financial identification, announced demand is therefore not a sufficient exposure measure: contractual status, legal-entity matching, denominator scope, and public timing all matter.

The central question is whether the volatility response to this joint pressure is super-additive. If climate stress and AI-related load merely contribute independent risks, a conventional additive variance specification is sufficient. If, however, contracted data-center load makes a utility more sensitive to heat-related grid stress—or if heat stress makes the market update the value of incremental AI load more abruptly—then the interaction should enter conditional variance directly. A second possibility is that the interaction operates through state dynamics: dual stress may raise the probability of entering a high-volatility regime or lengthen the time spent there. These are distinct economic claims and require distinct estimands.

This distinction matters because much of the existing empirical evidence is compatible with several competing interpretations. A positive association between climate news and volatility may reflect a common energy-price shock. A positive association between AI exposure and volatility may be little more than a technology, growth, or capital-expenditure factor. A post-2022 interaction between two trending indices may capture a period effect rather than a mechanism. The empirical design therefore uses utility-level, physically interpretable quantities. Climate exposure is tied to the service territory. AI exposure is tied to dated, operational or contractually binding customer-load megawatts. The denominator is the matched utility system peak load from EIA Form 861. Both components are lagged and standardized using expanding information sets.

The design is narrower than a general study of AI adoption. It does not infer AI use from a technology index, corporate text, or market classification. It studies operating or binding data-center electricity demand because that channel can be stated in physical units, matched to a utility system, and dated to the public information set. The audit also shows why the label matters: one qualifying load sequence, Otter Tail, predominantly reflects cryptocurrency hosting. The empirical estimand is therefore data-center power-demand exposure, with AI infrastructure as a motivating application rather than a verified treatment for every issuer.

The paper makes three contributions. First, it turns scattered load disclosures into a reproducible point-in-time measurement chain linking the parent security, operating utility, load object, EIA denominator vintage, weather territory, and public-information date. Second, it documents a concrete identification result: treating pre-disclosure silence as zero and multiplying uncentered regressors can make a climate-by-load interaction appear better supported than it is. Table 4A consolidates the full measurement-to-estimate path and distinguishes the OTTR label sensitivity from the common-positive and centering decisions. Third, the paper implements an advance rule that converts a failed first-stage interaction test into a disciplined stopping decision. Regime switching is reserved for a future sample that adds independent exposure variation, rather than used to rescue a weak single-state result.

The final broad data-center register contains ten listed parents: AEE, AEP, AES, CMS, D, EXC, IDA, MDU, OTTR, and PPL. The audit preserves the original filings, entity mappings, denominator vintages, weather records, market data, financial-control provenance, and corrections. The daily public-information panel contains 11,541 complete issuer-days. Requiring positive disclosed exposure for all ten parents produces a common window from 13 February through 31 December 2025, with 212 portfolio observations across nine variance-recursion segments.

The M1 estimates do not provide detectable support for the proposed amplification mechanism, but they do deliver an informative result about the design. In the legacy zero-coded panel, the uncentered interaction estimate is 0.143; excluding OTTR, whose tagged load is predominantly cryptocurrency hosting, reduces it to 0.060. These are label and coding sensitivities, not a sequential common-positive estimate. Reclassifying pre-disclosure zeros as unknown reduces the analysis to the common-positive window, where the uncentered interaction reaches -2.000 at its imposed bound. Centering the main effects produces an interior estimate of -0.118 , both information criteria select the no-interaction model, and a conditional parametric bootstrap yields $p=0.324$. Because M1 fails the advance condition, H3 remains untested and the paper does not estimate the Markov-switching extensions; this preserves a clean benchmark that can be updated when additional public disclosure history becomes available.

Section 2 develops the economic mechanisms. Section 3 documents the institutional setting and audited measurement. Section 4 states the directional hypotheses. Section 5 presents the staged estimation and stopping rule. Section 6 explains the identification limits. Section 7 reports data readiness and M1 results. Section 8 concludes.

2. Related Literature and Economic Mechanisms

2.1 Climate risk and financial volatility

Climate risk can affect asset prices through expected cash flows, financing costs, insurance and adaptation expenditures, policy exposure, and discount-rate variation. Evidence on carbon risk shows that investors distinguish firms by emissions and carbon intensity, while broader climate-finance research emphasizes the difficulty of converting long-horizon hazards into dated market innovations (Bolton and Kacperczyk, 2021; Giglio, Kelly, and Stroebel, 2021). A related measurement-layer problem appears when an issuer-level label compresses variation that remains visible at the underlying security or operating-entity level. In Chinese corporate bonds, public climate signals remain detectable in security-level pricing even when issuer credit ratings compress them (Yang, 2026b). That result motivates the present decision not to assign a parent-company climate or AI label mechanically: exposure is instead mapped to the regulated operating utility and its service territory. Physical risk is especially relevant for utilities because temperature and severe weather affect demand, operating costs, reliability, and investment requirements within geographically defined service territories.

Granular temperature evidence strengthens the case for location-specific measurement. Addoum, Ng, and Ortiz-Bobea (2023) link establishment-level temperatures to quarterly earnings and find economically meaningful temperature sensitivity across a broad set of industries. Acharya et al. (2022) exploit regional heat exposure across municipal bonds, corporate bonds, and equities, showing why headquarters-level or national averages can conceal the relevant asset location. These studies motivate the service-territory

mapping used here, but they do not test whether a newly expanding class of electricity demand changes the volatility transmission of heat stress.

For volatility measurement, the relevant insight is that climate information need not have a constant effect. A heat anomaly can be economically minor when capacity margins are ample and material when the grid is tight. The same event can also be priced differently across firms because service territories, generation mixes, and regulatory recovery mechanisms differ. This motivates a conditional-variance model with issuer-specific exposure rather than a national climate-news index alone.

2.2 AI adoption, correlated decisions, and market instability

AI can reduce idiosyncratic forecast error while increasing commonality in beliefs and actions. The distinction between forecasting performance and market stability is central: a model can be individually useful yet systemically destabilizing if many agents rely on similar inputs, architectures, or recommendations. Existing work discusses algorithmic herding, performative feedback, and cognitive dependency, but direct identification is difficult because adoption is often latent and because measures of AI exposure overlap with technology-sector composition (Yang, 2026a).

Recent theory makes the nonlinearity explicit. Meng and Chen (2026) model performative prediction, correlated algorithmic signals, and cognitive dependency as mutually reinforcing channels, implying a convex relationship between adoption and systemic coupling. Laboratory-style evidence using language-model agents is more ambiguous: Hansen and Lee (2025) find that AI agents can rely more heavily on private information than human subjects in some settings, yet can also herd when profit-maximizing incentives make such behavior optimal. Taken together, these studies provide mechanism language rather than a ready-made field measure. The present paper therefore does not infer trading-system adoption; it studies a separate, observable infrastructure channel through which AI expansion can alter the cash-flow sensitivity of utilities.

Electricity demand supplies a more concrete adoption margin. A binding large-load agreement reflects infrastructure commitment and expected consumption rather than generic AI enthusiasm. It can alter the utility investment path, load forecast, financing need, and regulatory bargain before the facility is fully operational. At the same time, contracted megawatts are not equivalent to realized consumption. The empirical measure therefore distinguishes operational load, binding approved load, and requested pipeline capacity, and treats disclosure dates as the beginning of the public information set.

2.3 Conditional volatility and regimes

ARCH and GARCH models formalize volatility clustering through the conditional second moment (Engle, 1982; Bollerslev, 1986). The GJR specification allows negative innovations to have a different effect from positive innovations of the same magnitude (Glosten, Jagannathan, and Runkle, 1993). This asymmetry is essential here because an apparent climate-by-AI interaction could otherwise be a relabeled leverage effect. Mixed-frequency extensions show how slow-moving macroeconomic variables can enter longer-run volatility components (Engle, Ghysels, and Sohn, 2013), while Markov-switching models allow parameters or transition probabilities to vary across latent states (Hamilton, 1989; Gray, 1996).

The paper uses these tools sequentially. A single-state GJR-GARCH-X is the first estimable benchmark. A two-state model is considered only if the data support the interaction and the simpler model

passes diagnostics. The regime framework then asks whether dual stress changes volatility within a state, changes the probability of leaving the low state, or changes the probability of remaining in the high state. This sequencing protects the economic interaction from being manufactured by an unstable latent-state fit.

2.4 The unresolved intersection

The three literatures leave a specific empirical intersection unresolved. Climate-finance studies generally treat electricity demand and grid capacity as part of the background through which weather affects firms. AI-finance studies generally measure adoption through text, patents, occupational exposure, or technology-sector affiliation. Volatility studies can accommodate nonlinear states but cannot identify an economic interaction when the underlying exposures are not separately measured. Consequently, finding that both climate and AI variables predict volatility would not establish that they reinforce each other.

The paper closes this gap by defining an estimand before selecting a nonlinear model. The numerator of AI exposure is a dated physical commitment; the denominator is the capacity scale of the utility serving that load; climate stress is local to the operating territory; and the interaction is constructed only after each component passes an independent audit. The Markov state is then used to decompose the interaction, not to rescue it. This measurement-before-regime ordering is the paper's central methodological commitment.

3. Institutional Setting, Sample, and Measurement

3.1 Sample construction

The study begins with U.S. listed electric-utility parents from 1 January 2021 through 31 December 2025. The original thirteen-name tariff register remains part of the provenance record. The empirical portfolio is instead the ten-parent disclosure-positive register—AEE, AEP, AES, CMS, D, EXC, IDA, MDU, OTTR, and PPL—formed by applying the frozen operating-or-binding customer-load rule. This is a retrospectively audited research sample, not an ex ante investable index.

Layer	Definition	Advance rule
Returns	Daily adjusted-close log returns; portfolio and issuer panels	At least 1,000 valid returns per included issuer and at least 90% common-date coverage
Climate	Lagged service-territory cooling-load stress	No revised or future information; representative station and territory mapping audited
Data-center load	Operating or binding customer-load MW divided by matched system peak MW	At least 10 issuers with four consecutive quarterly observations
Joint support	Audited C and A with verified zero-versus-unknown status	At least 50 complete observations in every central 2×2 cell
Model	Single-state GJR-GARCH-X before any regime model	Stable optimum and acceptable residual diagnostics

3.2 Returns and market controls

The dependent variable is 100 times the daily log total return. The principal series is an equal-weighted portfolio of the ten eligible parents, rebalanced at the start of each quarter; weights drift with total returns within the quarter. Market controls are the SPY return, XLU return, VIX log return, and

natural-gas return. Lag-available firm controls are portfolio-weighted noncurrent debt/assets, cash/assets, quarterly capital expenditure/assets, and regulatory assets/assets. AES regulatory assets remain an annual disclosure carried as the latest public observation with its source period and age preserved.

3.3 Physical climate stress

The primary climate variable is heat or cooling-load stress derived from daily maximum and minimum temperature at a representative service-area station in GHCN-Daily (National Centers for Environmental Information, 2026). Daily average temperature is converted into cooling degree days relative to 65°F. The anomaly is computed against a 1991–2020 calendar-day baseline and scaled by the historical monthly standard deviation. Heat is primary because it is the physical channel that directly overlaps with incremental data-center cooling demand and therefore gives the climate-by-load interaction a single interpretable mechanism. Cold snaps, storms, disasters, and outage megawatts can affect load and reliability through different signs, durations, and operational bottlenecks; combining them into one index would mix mechanisms and add timing noise. They remain pre-specified complementary physical-risk tests. For AES Ohio, the representative station is Dayton International Airport (GHCN-Daily USW00093815).

3.4 Data-center power-demand exposure

For utility i in quarter q , data-center power-demand exposure is $A_{i,q} = MW_{i,q} / \text{Peak}_{i,q}$. The numerator includes operational or contractually binding customer-load megawatts that can be assigned to the matched utility; requested or speculative pipeline megawatts and generator-side power-purchase agreements remain separate. The denominator is the matched utility system peak from the most recent Form EIA-861 vintage publicly available at the information cutoff (U.S. Energy Information Administration, 2026). Exposure begins after the disclosure date. This variable measures disclosed data-center commitment relative to system scale; it is not realized electricity use or a direct measure of firm-level AI adoption.

Evidence family	Numerator object	Treatment
A: utility customer load	Operational or binding data-center customer-load MW	Primary exposure; divide by matched utility system peak load
A-pipeline	Requested, queued, or nonbinding load MW	Separate robustness measure; never combined with committed MW
B: generator PPA	Generation output or PPA MW linked to AI/data-center counterparty	Separate event/robustness family; no utility-load denominator
Tariff/regulatory event	Large-load tariff, docket, or order without customer MW	Institutional evidence only; not relabeled as AI adoption

3.5 The interaction and no-lookahead timing

The first data build standardized climate and load variables using expanding information sets and coded exposure as zero before the first qualifying disclosure. The audit shows that this zero means “no qualifying disclosure observed,” not verified absence. The primary M1 sample therefore treats those dates as unknown and retains only positive disclosed exposure. For numerical identification, C and A are demeaned in the fixed 212-day portfolio sample and the interaction $(C - \bar{C})(A - \bar{A})$ is divided by its sample standard deviation. Main effects remain in the variance equation. This centering convention was fixed

before the final profile and bootstrap calculations and is reported because GARCH-X persistence restrictions can otherwise depend on arbitrary covariate origins.

3.6 AES Ohio as a measurement example

AES Ohio illustrates the distinction between a source-qualified observation and a complete exposure series. The listed security is AES Corporation. Its regulated Ohio operating utility is The Dayton Power and Light Company, doing business as AES Ohio, and the corresponding EIA Form 861 utility identifier is 4922. The 31 October 2024 SEC-filed earnings release states that agreements for an additional 900 MW brought 2024 year-to-date data-center load growth at AES Ohio to 2.1 GW (AES Corporation, 2024). Because the disclosure identifies the utility, customer class, contractual status, quantity, and publication date, 2.1 GW enters family A from that date forward.

The observation is nevertheless handled conservatively. It is not assigned to quarters preceding the filing. A later 310 MW statewide retail-supply item is recorded separately because it describes a different object and cannot be added mechanically to the utility load total. The EIA denominator uses only Dayton Power & Light's reported system peak, not consolidated AES generation or non-Ohio operations. Daily climate exposure is represented by Dayton International Airport station USW00093815. This chain—listed parent, legal utility, load object, peak-load denominator, weather territory, and public date—is required for every load-side observation.

3.7 Network capacity and regulatory confounds

Data-center siting is not random. Developers may select utilities with spare reserve margin, favorable generation mix, uncongested transmission, faster interconnection processing, or a regulatory regime that permits timely cost recovery. Those same characteristics can affect heat-event resilience, outage exposure, capital spending, and the speed at which financial information is incorporated. Thus, residualizing AI exposure on size, capital expenditure, leverage, and broad region does not remove grid-level confounding. A positive climate-by-load coefficient could otherwise reflect pre-existing network quality or regulatory maturity.

Available firm controls do not remove non-random data-center siting. Reserve margin, generation mix, transmission constraints, interconnection processing, and regulatory cost recovery may affect both load placement and heat resilience. Coverage of these network variables is not sufficient for a causal design in the current sample. Accordingly, the M1 estimates are conditional associations and the failure to reject the null is not evidence that the underlying physical mechanism is absent.

3.8 Exposure concentration and leave-one-out diagnostics

The predeclared concentration rule remains relevant to any future expansion: no broad cross-utility claim should rest on one unusually complete issuer. In the present version, leave-one-issuer-out regime diagnostics are not activated because M1 fails before M2. The OTTR exclusion is nevertheless reported because its audited load is predominantly cryptocurrency hosting rather than verified AI demand.

Excluding OTTR from the legacy full-history portfolio reduces the uncentered interaction estimate from 0.143 to 0.060. Restricting the same legacy specification to the 212 common-positive dates reverses

the sign and creates boundary problems. These comparisons motivated the unknown-aware classification and centered profile analysis; they are sensitivity diagnostics rather than estimates selected for their sign.

3.9 Audited data-foundation checkpoint

Ten issuers satisfy the broad data-center load-side four-quarter evidence rule: MDU, AEP, AES, AEE, EXC, D, IDA, PPL, CMS, and OTTR. Disclosure entry dates range from 9 January 2023 for OTTR to 12 February 2025 for Dominion. The source register records issuer filings and regulatory materials, including the AEP Ohio tariff filing, Ameren and Exelon investor materials, Dominion slides, and MDU annual reports (AEP Ohio, 2024; Ameren Corporation, 2024; Dominion Energy, 2025; Exelon Corporation, 2024; MDU Resources Group, 2024, 2025). G1–G3 therefore pass for returns, broad load evidence, and climate coverage. The legacy 2×2 grid has a minimum cell of 555 observations, but that count uses pre-disclosure zeros. Once those observations are classified as unknown, it cannot serve as a verified low-exposure control grid; confirmatory G4 support is therefore unresolved. Table 5 reports the nonduplicated gate summary.

The denominator and climate mappings use the same operating utility and point-in-time public vintages. For Exelon, the qualifying 2024Q4–2025Q3 window uses ComEd's 2023 EIA vintage peak of 22,467 MW; the 2024 vintage peak of 21,560 MW enters only in 2025Q4. AEP Ohio uses 7,914, 8,104, and 9,014 MW for the successive 2022–2024 vintages after their release dates. NOAA temperature records are converted from tenths of degrees Celsius before computing Fahrenheit temperatures, cooling degree days, normals, and heat stress; nonblank quality flags are excluded.

4. Hypotheses

H1 — variance amplification. The climate-by-load interaction increases conditional variance, with the largest effect expected in the high-volatility state.

H2a — **state entry**. Dual stress lowers the probability of remaining in the low-volatility state.

H2b — **state persistence**. Dual stress raises the probability of remaining in the high-volatility state and increases expected high-state duration.

H3 — **exposure gradient**. The marginal response to climate stress is larger for the High-A portfolio than for the Low-A portfolio.

H1 is tested at M1 and reported regardless of sign. H2a, H2b, and H3 remain directional design commitments but are not estimated because the advance rule stops the sequence before M2. No transition or heterogeneity result is implied.

4.1 Mechanism-specific observable implications

The cash-flow channel predicts that the interaction is strongest when heat stress coincides with high committed-load intensity and when cost recovery is delayed or uncertain. The information channel predicts larger effects around disclosures that revise contracted load or capital requirements, but not around generic AI news unrelated to the utility territory. The system-pressure channel predicts that the interaction is more visible in the upper tail of conditional variance and in low-to-high transition probabilities than in the conditional mean.

These implications also define informative failures. A climate effect with no High-A gradient supports a general weather-risk channel but not the proposed interaction. An AI-load main effect that survives during low climate stress may indicate financing or growth uncertainty rather than dual pressure. A transition effect without stable state separation is treated as model failure. A significant interaction driven only by future exposure leads or by pipeline requests is classified as timing or measurement contamination.

5. Empirical Framework

5.1 Single-state benchmark

The M1 benchmark is a standardized Student-t GARCH-X model. Returns follow $r_t = \mu + \phi r_{t-1} + b'X_t + \varepsilon_t$, with $\varepsilon_t = \sigma_t z_t$. After the boundary audit sets the empirically negligible GJR leverage term to zero, the reported variance recursion is $\sigma_t^2 = (\omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2) \exp[\lambda^c(C_{t-1} - \bar{C}) + \lambda^A(A_{t-1} - \bar{A}) + \lambda^{cA}Q_{t-1}]$, where Q is the centered product divided by its sample standard deviation. The exponential term preserves positivity; λ^{cA} is the log-variance scaling associated with a one-standard-deviation change in the centered interaction coordinate.

5.2 Markov-switching extension

The planned two-state model remains a contingent extension. If M1 had passed, low- and high-volatility states would have been ordered by conditional variance and transition probabilities would have included the same centered exposure coordinates. Because M1 provides no incremental interaction support, estimating this more heavily parameterized model would violate the advance rule.

No transition coefficient, marginal effect, state occupancy, or expected duration is reported. This is a substantive stopping decision: latent states are not used to manufacture an interaction that the simpler observed-state benchmark does not support.

5.3 Estimation, inference, and diagnostics

M1 is estimated by conditional maximum likelihood from dispersed starts. The 212 observations form nine segments because a new segment is declared whenever one or more candidate trading dates are excluded by the common-positive or public-information rule. At each boundary, the conditional-variance recursion is reset to the same variance of the first 60 OLS residuals; σ^2 is not propagated across the omitted dates. The mean equation retains the last observed portfolio return as r_{t-1} , so the reset applies only to the latent variance state. The bootstrap uses the identical fixed reset vector and re-estimates the initialization variance in every draw. The boundary audit compares nonnegative free-GJR and gamma-fixed-zero specifications under three segment-initialization multipliers. An additional fixed-zero warm start corrected a local optimum in the null model. The centered specification then fixes gamma at zero, uses 20 dispersed starts for the interaction model, and verifies the Student-t likelihood independently.

Inference uses a 499-replication parametric bootstrap under the centered no-interaction model, conditional on the observed regressor matrix. Each simulated sample recomputes the first-60-observation OLS residual variance used for segment initialization and is fitted from two null and three interaction starts. All 499 replications produce valid solutions; the observed LR statistic is 1.638 and the plus-one

bootstrap p-value is 0.324. A descriptive profile-likelihood interval is also reported, but it is not treated as a substitute for the bootstrap. The fixed-design bootstrap conditions on the observed regressors and segment pattern; it is not a causal test or a regime-occupancy power analysis.

Stage	Specification	Question	Advance condition
M0	Event-window realized-volatility contrasts	Are sign and timing consistent with the physical mechanism?	Descriptive validation only
M1	Single-state GJR-GARCH-X	Is there stable additive or interactive variance scaling?	Convergence and residual diagnostics
M2	Two-state model, constant transitions	Is amplification concentrated in a high-volatility state?	Separated, occupied, stable states
M3	Time-varying transitions	Does dual stress affect entry and persistence?	Effective transition variation and bootstrap feasibility
M4	High-A/Low-A and bounded robustness families	Is the response exposure-specific and definition-robust?	Report regardless of sign

5.4 Finite-sample feasibility and parameter budget

The realized parameter budget is 16 free parameters in the centered no-interaction model and 17 in the interaction model, including ten mean parameters, the Student-t tail parameter, variance dynamics, and exposure slopes. The apparent daily panel is much larger than the identifying time series: the unknown-aware portfolio contains 212 observations across nine segments, while exposure changes occur only at disclosure and denominator-vintage dates. This distinction is central to feasibility.

Rather than report post hoc “observed power,” which is mechanically tied to the realized test statistic, the paper reports a bootstrap-dispersion sensitivity calculation. Under the fitted null, the standard deviation of the 499 unrestricted interaction estimates is 0.101. A local-normal approximation therefore gives an 80%-power minimum detectable absolute coefficient of $(1.96+0.842) \times 0.101 = 0.284$ at a two-sided 5% level. Because λ^{CA} is a log-variance multiplier, $|0.284|$ corresponds to approximately a 33% variance change per one-standard-deviation movement in the centered product. This threshold is economically material in the selected no-interaction model: its fitted conditional variance rises by about 41% from the sample 25th percentile (0.192) to the 75th percentile (0.272), and by about 76% from the 5th percentile (0.183) to the 95th percentile (0.322). It also exceeds the 95th percentile of within-segment one-day absolute log-variance changes (0.079, approximately 8%). A 33% interaction multiplier is therefore slightly smaller than the model’s interquartile cross-date variance spread but much larger than an ordinary one-day variance movement. The current design is informative about large amplification and remains weak for moderate effects. This is a diagnostic approximation conditional on the observed design, not a guarantee of power against every nonlinear alternative.

The originally planned M3 Monte Carlo exercise is not run because the sequence stops at M1. Its state-occupancy criteria cannot validate a regime model after the prerequisite observed-state interaction has failed. Any future return to M2/M3 requires a longer positive-exposure history or additional independently audited issuers, followed by a newly frozen feasibility design.

6. Identification and Reporting Design

Physical load intensity reduces overlap with a traded technology factor, but it does not create exogenous treatment. The ten-parent register measures data-center commitments, and OTTR is predominantly tied to cryptocurrency hosting. Market, sector, energy, and firm controls are included, while network capacity and regulatory confounds remain incompletely observed. The estimates are therefore descriptive conditional-volatility evidence.

Common-trend and disclosure problems are addressed by retaining the legacy coding only as a sensitivity case and treating pre-disclosure observations as unknown in the primary M1 sample. This choice reduces the usable portfolio history to 212 dates. The paper reports that loss of support rather than converting non-disclosure into a low-exposure control group.

The original central-cell count is not sufficient for confirmatory interaction inference because the low-A cells partly rely on unverified zeros. Before recentering in the unknown-aware sample, the product is highly aligned with the heat main effect: $\text{Corr}(C, C \times A) = 0.979$ and the interaction VIF is 29.4. After centering, $\text{Corr}(C - \bar{C}, Q) = -0.127$, $\text{Corr}(A - \bar{A}, Q) = -0.249$, and the interaction VIF falls to 1.09. Centering therefore removes the mechanical coordinate collinearity and the artificial coefficient-bound result, but it does not create new exposure dates or independent economic variation. Table 4A summarizes this sequence and shows that the 0.060 estimate is an OTTR-exclusion sensitivity, not an intermediate common-positive estimate. The remaining limitation is power and effective exposure support, not conventional linear multicollinearity in the centered regressors.

Table 4A. Measurement decisions and the interaction estimate

Stage	Exposure/sample decision	N	λ_{CA}	ΔAIC	VIF	Diagnostic interpretation
Legacy panel	Pre-disclosure silence coded zero; all ten issuers	1,118	0.143	-2.736	1.93	Pilot AIC favors interaction; no calibrated inference
Label sensitivity	Legacy coding; OTTR excluded	1,119	0.060	+1.097	1.59	AIC reverses; crypto-hosting label matters
Unknown-aware	All ten issuers positive; uncentered $C \times A$	212	-2.000	-0.037	29.39	Coefficient bound; $\text{Corr}(C, C \times A) = 0.979$
Final centered M1	Unknown-aware; centered and scaled product	212	-0.118	+0.362	1.09	Interior; bootstrap $p = 0.324$

Note. The rows are diagnostic specifications, not a sequence of confirmatory hypothesis tests. The -2.000 estimate is at the imposed coefficient bound and is not interpreted economically.

Model adequacy is evaluated across initialization multipliers, leverage restrictions, starts, information criteria, profile likelihood, and the conditional bootstrap. The no-interaction model is preferred under AIC and BIC in every boundary/initialization configuration. These diagnostics support the stop decision, while their fixed-design and short-sample limitations prevent a structural or causal conclusion.

The preserved data can support future work on heat-sensitive utility volatility, data-center load disclosure, denominator vintages, and event timing. A future interaction test should extend the post-disclosure history and obtain independently varying exposures. Storm/outage measures, generator PPAs, GARCH-MIDAS, Wasserstein diagnostics, and Markov switching remain separate extensions rather than specifications searched after observing M1.

7. Data Readiness and M1 Results

7.1 Audited measurement sample

The data audit yields ten broad data-center exposure sequences and 11,541 issuer-days with market, climate, and latest-public financial controls. Of the 12,540 issuer-day records in the underlying panel, 8,685 pre-disclosure zero codes are reclassified as unknown and 3,855 observations retain positive disclosed commitments. The requirement that all ten portfolio constituents have positive disclosed exposure and complete controls leaves 212 portfolio observations across nine variance-recursion segments from 13 February to 31 December 2025. Table 5 records how this audit changes the advance decision; Table 7 preserves issuer-level descriptive heterogeneity without treating it as an H3 test.

Table 5. Data-readiness and advance status

Gate	Observed evidence	Status
G1 Returns	Ten-issuer portfolio returns available; 212 unknown-aware common-positive dates	PASS for M1
G2 Load evidence	Ten broad data-center sequences; OTTR mainly crypto hosting; not ten AI-specific measures	PASS only for broad DC load
G3 Heat stress	Matched NOAA service-territory series; corrected 0.1°C units	PASS
G4 Joint support	Legacy minimum cell 555 relies partly on unverified pre-disclosure zeros	UNRESOLVED
G5 M1 Regime advance	Centered bootstrap LR=1.638, p=0.324; AIC/BIC prefer null All prerequisites required	FAIL STOP

Note. G2 validates a broad data-center load measure only. It does not certify ten AI-specific adoption or AI-load measures.

7.2 Single-state interaction evidence

The uncentered legacy estimates are sensitive to sample definition and coefficient bounds. Under the primary unknown-aware sample, centering C and A and scaling their product to unit standard deviation yields an interior interaction estimate of -0.118 . The interaction lowers negative log likelihood by 0.819, but the additional parameter raises AIC from 323.171 to 323.533 and BIC from 376.876 to 380.595. Thus both criteria select the no-interaction specification in this sample. The bootstrap p-value of 0.324 fails to reject the no-interaction model; it does not establish that the population interaction is zero.

Table 6. Centered Student-t GARCH-X comparison

Specification	NLL	Parameter s	AIC	BIC	λ_{CA}	Bootstrap result
No interaction	145.585	16	323.171	376.876	—	Reference
Centered interaction	144.767	17	323.533	380.595	-0.118	LR=1.638; p=0.324

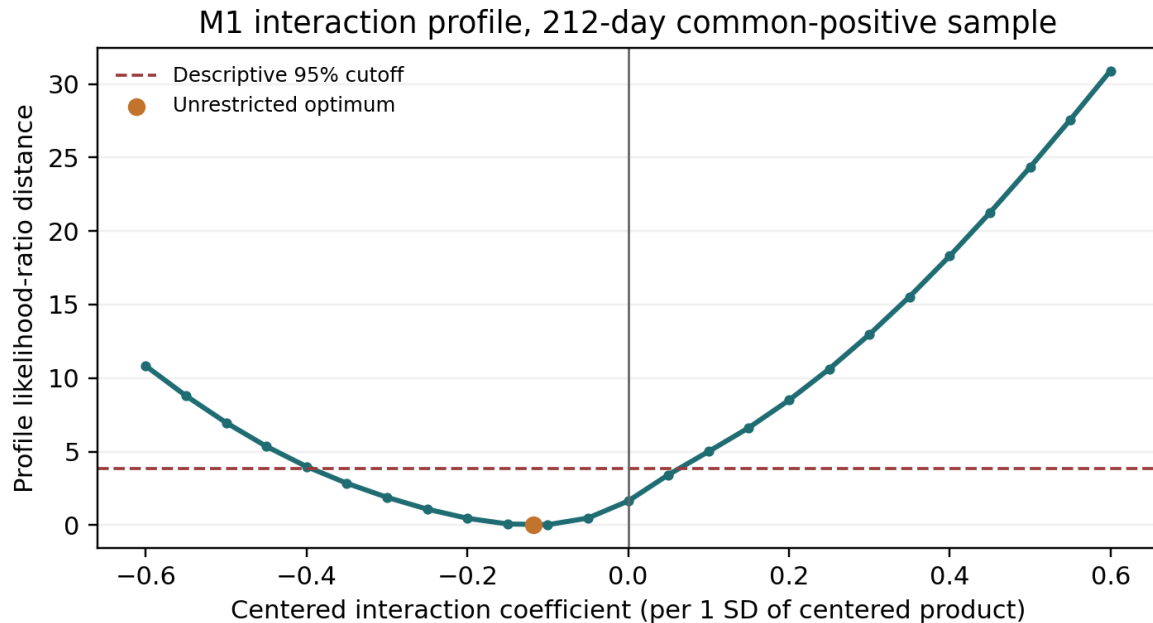


Figure 1. Profile likelihood for the centered interaction coefficient. The dashed line is the descriptive one-parameter 95% cutoff; the interpolated interval is $[-0.395, 0.064]$ and includes zero.

7.3 Interpretation and stopping rule

The estimates do not establish amplification, absorption, or a causal null. They show that the audited 2021–2025 sample lacks detectable incremental support for the proposed interaction after disclosure uncertainty and coordinate dependence are handled. This failure to reject is scientifically useful because it locates the present bottleneck: effective exposure variation, rather than conventional centered-regressor collinearity or optimizer failure. The approximate 80%-power threshold of $|0.284|$ indicates that economically smaller interactions remain unresolved. In accordance with the predeclared sequence, G5 is marked FAIL and M2/M3 are not estimated. Every source decision, denominator vintage, coding change, and optimization attempt is retained, allowing a longer post-disclosure panel to revisit the same hypothesis without redefining the evidence after seeing the result.

7.4 Issuer-level descriptive coverage

The portfolio test deliberately avoids treating issuer-days as independent replications of a common quarterly exposure. Table 7 nevertheless reports the cross-sectional scale hidden by aggregation. The wide range is economically plausible under the frozen definition because the numerator is disclosed operating or binding customer load and the denominator is the matched regulated utility's system peak—not the consolidated parent's size or current realized data-center consumption. At the low end, EXC uses 500 MW relative to Commonwealth Edison peaks of 22,467 MW through 2025Q3 and 21,560 MW in 2025Q4. At the high end, Dominion uses 14,000 MW relative to Virginia Electric & Power peaks of 17,957–18,409 MW; PPL uses 5,000 MW relative to 6,898–7,303 MW; and AES uses 2,100 MW relative to Dayton Power & Light peaks of 3,241–3,280 MW. MDU's range reflects a dated numerator change from 180 MW to 530 MW against a much smaller 652–675 MW utility peak. These ratios measure

prospective committed-load burden relative to system scale and can therefore be large without implying that the load was already operating at that share. The numerator/denominator pairs were rechecked against the point-in-time source register; no denominator remapping is introduced in this revision. The figures document heterogeneity in levels but do not estimate H3: disclosure windows differ, load objects are not uniformly AI-specific, and no issuer-level significance tests are used to revive a failed portfolio gate.

Table 7. Issuer-level positive-disclosure coverage and exposure intensity

Issuer	First positive date	Positive days	Mean A (%)	A range (%)	Mean r (%)
AEE	2024-11-08	286	3.39	3.37–3.47	0.82
AEP	2024-05-14	410	7.34	6.66–7.58	0.87
AES	2024-11-01	291	64.64	64.02–64.79	2.02
CMS	2024-11-01	291	3.06	3.06–3.08	0.84
D	2025-02-13	222	77.45	76.05–77.96	1.01
EXC	2024-11-11	285	2.25	2.23–2.32	0.90
IDA	2023-12-29	503	3.87	3.69–3.92	0.82
MDU	2024-02-23	466	52.23	27.56–81.28	1.15
OTTR	2023-01-10	747	10.45	10.12–11.04	1.22
PPL	2024-08-05	354	71.35	68.47–72.48	0.84

Note. A is committed data-center customer-load MW divided by the matched utility peak-load denominator. “Positive days” counts disclosed-positive issuer-days through 31 December 2025. Mean |r| is descriptive and is not adjusted for market or firm controls.

8. Conclusion

This paper asks whether disclosed data-center electricity demand changes how local heat stress is translated into utility-stock volatility. The immediate answer is deliberately limited: the present common-positive sample does not detect incremental amplification. The durable contribution is that the study makes the test well defined. It provides a point-in-time chain linking parent securities, operating utilities, customer-load evidence, system-peak denominators, weather territories, returns, and lag-available financial controls. It also demonstrates that broad data-center exposure is not uniformly AI-specific, that pre-disclosure silence is not a verified zero, and that thousands of issuer-days can collapse to only 212 portfolio observations across nine variance-recursion segments once common information is required.

The M1 evidence fails to detect a climate-by-load amplification effect; it is not evidence that the effect is absent. H3 remains untested. The positive empirical lesson is that measurement choices determine whether an apparent interaction exists. Reclassifying disclosure silence as unknown changes the admissible sample, while centering reduces the interaction VIF from 29.4 to 1.09 and removes an artificial coefficient-bound result. AIC and BIC then select the no-interaction specification, the bootstrap does not reject it, and the approximate 80%-power threshold is $|0.284|$ log-variance units. That threshold is slightly below the fitted null model’s 41% interquartile cross-date variance spread, showing that the design can detect only economically large amplification with reasonable power. Stopping before regime switching is consequently part of the contribution rather than an unfinished analysis. The frozen evidence register and estimation archive create a reusable baseline: when longer disclosure histories or more independently varying audited exposures become available, the same gate can be rerun and the nonlinear extension can proceed without changing the original definitions. The paper therefore converts an inconclusive interaction estimate into a transparent measurement result, a falsifiable update rule, and a durable empirical asset.

Declarations

Conflict of interest. The author declares no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Ethics approval. Not applicable. The study uses publicly available market, regulatory, corporate-disclosure, energy-system, and weather data and involves no human participants.

Data availability. The frozen replication archive is `Climate_AI_GARCH_Reusable_Data_v0.21`, comprising core processed data and code, SEC evidence, NOAA/EIA source material, and SHA-256 manifests. M1 sensitivity and centered-profile supplements v0.22–v0.23 preserve every optimizer start, likelihood check, bootstrap draw, and reporting input. The SSRN v1.2 closeout package preserves the centered VIF calculation, segment map, minimum-detectable-effect and conditional-variance calibration, issuer-level numerator/denominator audit, citation crosswalk, and the exact upload manuscript. Public-source terms and redistributability remain those of the original providers.

Use of generative AI. Generative AI tools assisted language editing, source triage, code scaffolding, and bibliographic cross-checking. They did not generate or impute observations, decide which megawatt evidence qualified, recode failed observations to obtain a preferred sign, select specifications after viewing significance, or authorize advancement after a failed gate. The author manually reviewed source selection, variable definitions, estimation diagnostics, calculations, citations, and claims.

References

- Acharya, V. V., Johnson, T. C., Sundaresan, S. M., and Tomunen, T. (2022). Is physical climate risk priced? Evidence from regional variation in exposure to heat stress. NBER Working Paper No. 30445. <https://doi.org/10.3386/w30445>.
- Addoum, J. M., Ng, D. T., and Ortiz-Bobea, A. (2023). Temperature shocks and industry earnings news. *Journal of Financial Economics*, 150(1), 1–45. <https://doi.org/10.1016/j.jfineco.2023.07.002>.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), 307–327. [https://doi.org/10.1016/0304-4076\(86\)90063-1](https://doi.org/10.1016/0304-4076(86)90063-1).
- 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>.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4), 987–1007. <https://doi.org/10.2307/1912773>.
- Engle, R. F., Ghysels, E., and Sohn, B. (2013). Stock market volatility and macroeconomic fundamentals. *Review of Economics and Statistics*, 95(3), 776–797. https://doi.org/10.1162/REST_a_00300.
- 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>.

- Glosten, L. R., Jagannathan, R., and Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance*, 48(5), 1779–1801. <https://doi.org/10.1111/j.1540-6261.1993.tb05128.x>.
- Gray, S. F. (1996). Modeling the conditional distribution of interest rates as a regime-switching process. *Journal of Financial Economics*, 42(1), 27–62. [https://doi.org/10.1016/0304-405X\(96\)00875-6](https://doi.org/10.1016/0304-405X(96)00875-6).
- Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57(2), 357–384. <https://doi.org/10.2307/1912559>.
- Hansen, A. L., and Lee, S. J. (2025). Financial stability implications of generative AI: Taming the animal spirits. Finance and Economics Discussion Series 2025-090. Board of Governors of the Federal Reserve System. <https://doi.org/10.17016/FEDS.2025.090>.
- Meng, S., and Chen, X. (2026). Artificial intelligence and systemic risk: A unified model of performative prediction, algorithmic herding, and cognitive dependency in financial markets. arXiv:2604.03272. <https://doi.org/10.48550/arXiv.2604.03272>.
- Shehabi, A., Smith, S. J., Hubbard, A., Newkirk, A., Lei, N., Siddik, M. A. B., Holecek, B., Koomey, J. G., Masanet, E., and Sartor, D. (2024). 2024 United States Data Center Energy Usage Report. Lawrence Berkeley National Laboratory, LBNL-2001637. <https://doi.org/10.71468/P1WC7Q>.
- Yang, F. (2026a). Artificial intelligence and financial volatility: From model augmentation to market disruption—A systematic critical review. SSRN. <https://doi.org/10.2139/ssrn.7196800>.
- Yang, F. (2026b). Where public climate signals remain visible: Issuer compression and debt-rating sorting in China. *Journal of Financial Reporting and Accounting*, forthcoming.

Primary data and disclosure sources

- AES Corporation (2024). Third Quarter 2024 Earnings Release, filed 31 October 2024. U.S. Securities and Exchange Commission.
- U.S. Energy Information Administration (2026). Annual Electric Power Industry Report, Form EIA-861, 2001–present.
- National Centers for Environmental Information (2026). Global Historical Climatology Network Daily (GHCN-Daily).
- AEP Ohio (2024). AEP Ohio Files Plan to Secure Grid Resources for Data Centers, Protect Residential Customers, 13 May 2024; Public Utilities Commission of Ohio Case No. 24-0508-EL-ATA.
- Ameren Corporation (2024). Third Quarter 2024 Earnings Presentation, 7 November 2024.
- Dominion Energy, Inc. (2025). Fourth Quarter 2024 Earnings Call Slides, 12 February 2025.
- MDU Resources Group, Inc. (2024, 2025). Annual Reports on Form 10-K for fiscal years 2023 and 2024.
- Exelon Corporation (2024). Investor Meetings, November 2024, slide 16: Historical Data Center Load in Northern Illinois.