Peer reviewed Journals

Sci Chronice

RESEARCH

Accountability Signals After Large Service Outages and Reputation Recovery Across Customer Investor and Employee Audiences

Amitabh Sharma,[1] Rakesh Pradhan,[2] and Manish Kulshrestha[3]

[1]Department of Business Administration, Central University of Himachal Pradesh, Dharamshala Campus, Shahpur Road, Dharamshala 176215, Himachal Pradesh, India
[2]School of Commerce and Management Studies, Central University of Odisha, Sunabeda Campus, NAD Post Office, Koraput 763004, Odisha, India
[3]Department of Finance and Accounting, Dr. Harisingh Gour Vishwavidyalaya, University Road, Sagar 470003, Madhya Pradesh, India

Abstract

Large service outages convert internal operating failures into public tests of corporate reliability. When airline systems fail, payment networks stall, cloud platforms go offline, hospitals lose scheduling access, or delivery platforms stop routing orders, stakeholders observe not only the disruption but also how the firm explains responsibility, remediation, and prevention. This paper examines how accountability signals after major outages affect reputation recovery and market reactions. The empirical sample contains 1,472 service-outage events among 268 publicly traded firms from 2014 to 2024. The dataset combines outage telemetry, customer complaints, app and web activity, call-center wait proxies, employee text, press releases, executive statements, equity returns, option-implied volatility, analyst reports, and customer churn proxies. An accountability signal is measured by whether the firm identifies the failure mechanism, accepts responsibility, specifies customer remedies, provides restoration milestones, and commits to verifiable prevention. The analysis applies matched event studies, synthetic outage counterfactuals, accelerated recovery models, double machine learning, and causal forests. The results show that outages generate an average three-day abnormal return of minus 0.76%, but firms issuing high-specificity accountability signals experience 39.4% faster customer sentiment recovery and 2.18% lower implied-volatility persistence. Compensation-only responses reduce short-run complaint volume but do not restore reliability perceptions unless paired with technical explanation. Executive apologies are useful only when accompanied by operational milestones. The evidence indicates that accountability signals preserve reputation when they reduce uncertainty about recurrence, not when they merely express regret.

1. Introduction

Operational reliability is a public dimension of corporate reputation. Many firms sell continuity, access, routing, payments, reservations, delivery, account visibility, streaming availability, or service coordination. Customers do not observe the internal architecture that supports these services. They observe whether the service works, whether disruptions are acknowledged, whether remedies are provided, and whether the firm appears capable of preventing recurrence. Large outages therefore create a visible test of competence. A failure may begin as a technical interruption, staffing breakdown, vendor error, software release problem, capacity failure, or security precaution, but it becomes a corporate signal once the firm communicates about it.

This paper studies accountability signals after large-scale service outages. The empirical setting includes airlines, payment processors, banks, telecommunications providers, cloud platforms, streaming services, delivery platforms, health service networks, logistics firms, insurance portals, online brokerages, ride-hailing systems, travel platforms, and subscription software firms. The focal event is not a privacy controversy, supplier misconduct, workplace voice episode, endorser scandal, voluntary metric opacity, or ordinary customer-review decline. It is an operating interruption that prevents stakeholders from using a promised service. The firm's post-outage communication is then examined as a corporate signal whose credibility affects reputation recovery and market reaction.

The main argument is empirical and conditional. A service outage can damage reputation because it reveals fragility in a system that stakeholders expected to be reliable. The firm's response can reduce that damage when it clarifies what failed, who is responsible for recovery, how customers will be made whole, when restoration will occur, and what will change to prevent recurrence. A response that expresses regret but gives no mechanism leaves stakeholders uncertain about repeat failure. A response that offers compensation but no diagnosis may reduce anger while leaving competence doubts unresolved. A response that gives techni-

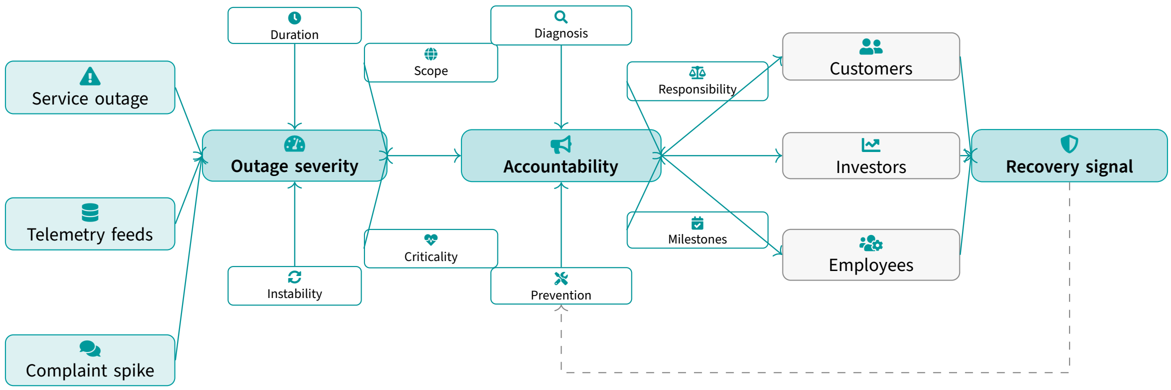

Figure 1. Accountability signal formation after a public service outage. Operational evidence is converted into severity, response specificity, stakeholder interpretation, and observed recovery.

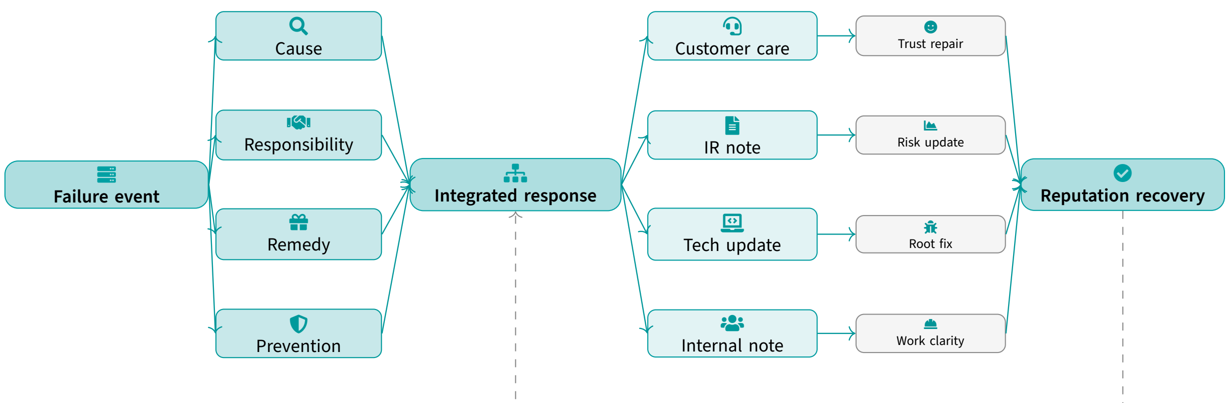

Figure 2. Response architecture across customers, investors, technical users, and employees. Recovery is strongest when the firm converts the outage into a coherent diagnosis-remedy-prevention sequence.

cal detail but ignores customer harm may help analysts yet fail with customers. The paper therefore separates apology, diagnosis, remedy, timeline, and prevention commitment.

The study uses 1,472 outage events among 268 publicly traded firms from 2014 to 2024. Events are included when the outage affects a material customer-facing service, is publicly observable, and can be timed to a specific date. The dataset integrates outage-monitoring feeds, firm status pages, app telemetry, customer complaint spikes, web traffic, press coverage, firm statements, executive social posts, earnings-call discussion, stock returns, option prices, analyst reports, and customer churn proxies [1]. The design observes the outage, the accountability signal, stakeholder reaction, and market reaction around the same event.

The paper treats accountability communication as an observable signal of managerial competence and stakeholder orientation. Murshed et al. (2023) [2] show that ethical corporate social responsibility and interactional justice can shape employee identification and job satisfaction through perceptions of fairness and relational treatment. This finding is relevant because outage response also has a fairness component: customers, employees, and investors evaluate whether the firm acknowledges harm, communicates respectfully, and provides remedies in a way that appears procedurally credible.

The empirical design differs from a simple apology study. Apologies are measured, but they are not assumed to be effective [3]. The analysis tests whether apologies work only when bundled with operational content [4]. The central measure is accountability specificity, which captures failure diagnosis, responsibility acceptance, restoration timeline, customer remedy, and prevention commitment. The study then estimates how accountability specificity affects customer sentiment recovery, complaint decay, app engagement, churn proxies, implied volatility, abnormal returns, analyst language, and employee morale.

The results show that the average outage produces a negative market reaction. The mean three-day abnormal return is minus 0.76%, and thirty-day implied volatility rises by 3.08%. Customer complaint volume rises by 64.2% in the first week after the outage, while negative customer sentiment rises by 41.7%. The average firm recovers customer sentiment in 6.8 weeks, but recovery time varies substantially. High-accountability

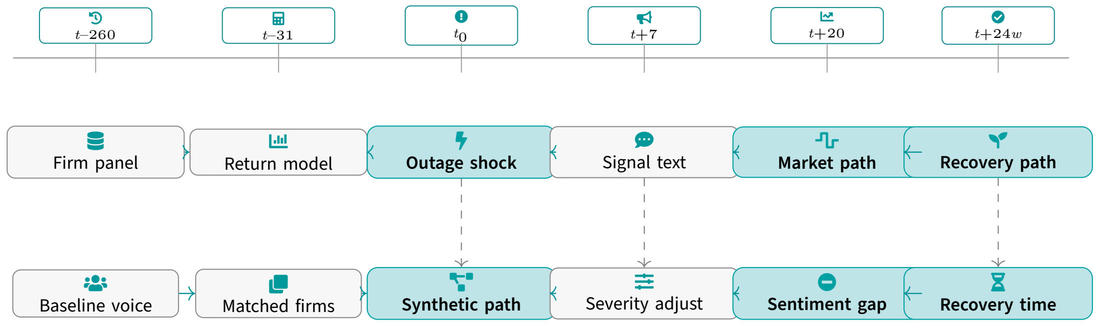

Figure 3. Event-study and synthetic-counterfactual timing. The design separates pre-event expected performance, the outage shock, response-window accountability, and post-event recovery paths.

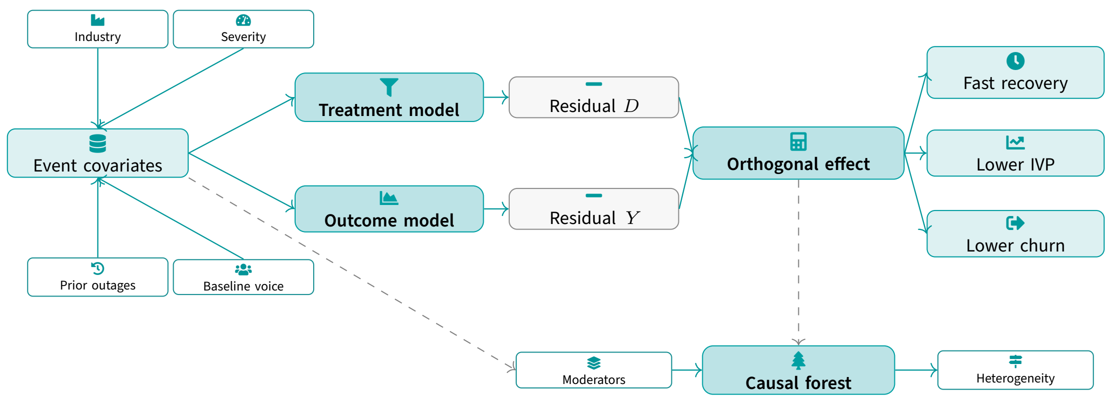

Figure 4. Double machine-learning and causal-forest validation. Severity-adjusted accountability is isolated from flexible nuisance structure before estimating average and heterogeneous recovery effects.

responses shorten median sentiment recovery time by 39.4%. They also reduce the persistence of option-implied volatility and analyst concern. Compensation-only responses reduce complaint volume in the first two weeks but do not restore reliability sentiment unless paired with failure explanation and prevention milestones.

The paper also finds that the identity of the communicator matters less than the content. Chief executive or founder apologies receive more media attention, but they do not consistently improve recovery when they lack operational detail. Technical status updates reduce analyst uncertainty but have limited customer effect when they ignore customer disruption. Joint messages that combine executive accountability, technical diagnosis, customer remedy, and prevention commitment produce the fastest recovery. The implication is that the accountability signal is multidimensional; symbolic humility is not a substitute for diagnostic clarity.

The effects differ across industries. Payment, brokerage, health-service, travel, and logistics outages generate stronger market penalties than entertainment outages because access reliability is more central to perceived service value. Cloud and enterprise software outages generate large analyst and option-market responses even when consumer sentiment is limited, because business customers and future service-level obligations are affected. Consumer app outages generate rapid complaint spikes but faster recovery when switching costs are high and communication is clear. Airline and travel platform outages generate persistent sentiment damage when customer remedies are vague or delayed.

Cao and Yan (2016) [5] show that the specificity of stakeholder-facing product claims can be positively related to firm performance, while excessive diversity in such claims can weaken performance. The present study uses this insight in a different context by examining whether specific accountability content after service failure restores stakeholder confidence more effectively than broad expressions of concern or scattered response messages.

The contribution is threefold. First, the paper identifies accountability specificity as a corporate signal after operating disruption. Second, it connects accountability signals to both reputation recovery and market reaction, rather than studying customer sentiment or stock prices separately. Third, it uses outage telemetry and machine learning to separate the severity of the failure from the quality of the response. This separation is crucial because severe outages often receive more detailed responses, which can obscure whether detailed responses help recovery. The empirical strategy uses matched comparisons, double machine learning,

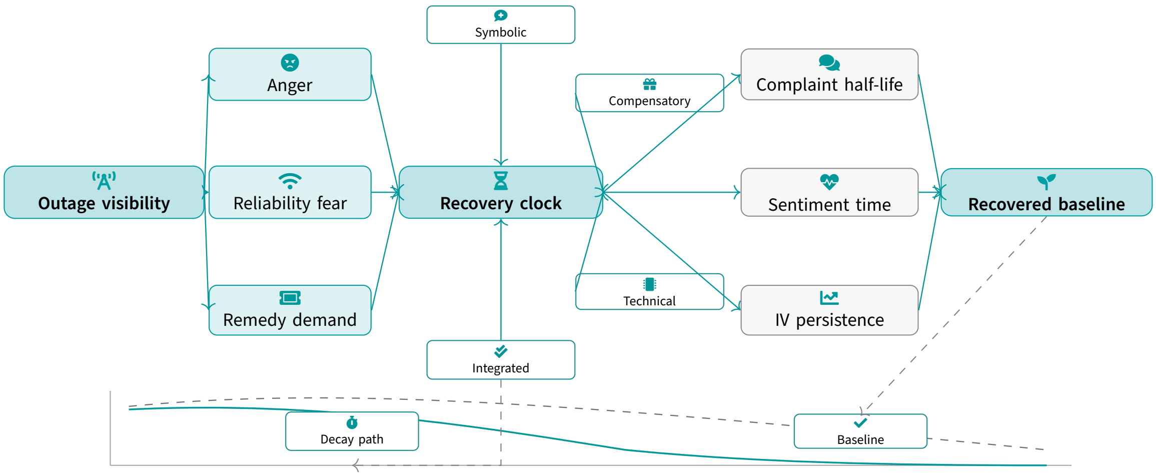

Figure 5. Recovery dynamics after the outage. Different response patterns shorten or prolong complaint decay, reliability-sentiment normalization, and option-market uncertainty.

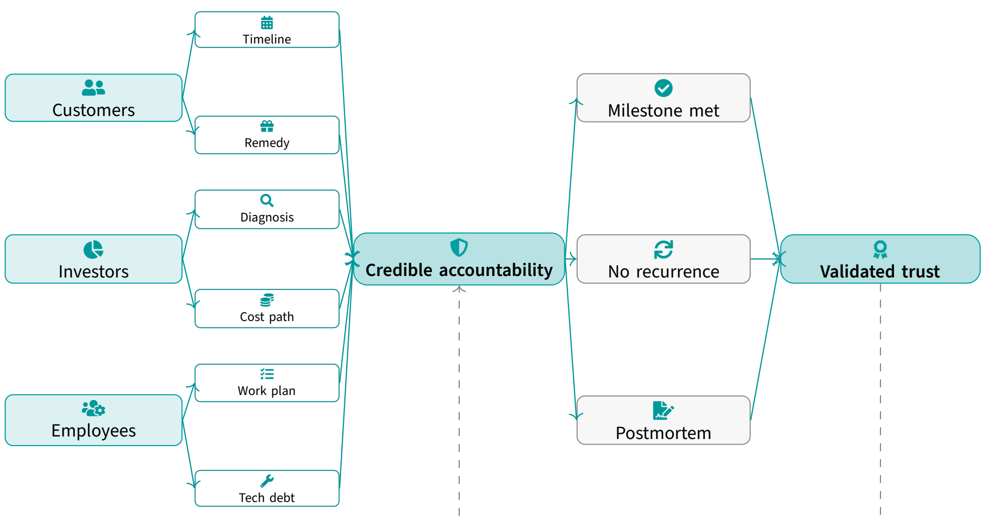

Figure 6. Audience-specific information needs and validation. Customers, investors, and employees require different parts of the response, but later operating evidence determines whether the signal becomes credible.

and causal forests to address this issue.

The study does not claim that communication can erase operational failure. In many outages, reputation recovery depends on actual service restoration and later recurrence. A clear statement without operational correction is fragile. The empirical results show that accountability signals are most effective when later recurrence is low. The paper therefore interprets accountability specificity as a credible signal when it is aligned with restoration and prevention, not as rhetoric alone.

2. Data and Event Formation

The sample begins with public service-outage events among U.S.-listed firms from 2014 to 2024. An outage enters the sample when a customer-facing system becomes materially unavailable, degraded, or unreliable for a defined period and when the outage is visible through customer complaints, monitoring feeds, platform status pages, regulatory notices, firm statements, or press coverage. Events are excluded when the main issue is a data breach, privacy controversy, labor action, supplier misconduct, financial fraud, product recall, or deliberate maintenance window. This screening keeps the empirical focus on operational service interruption.

The final event file includes 1,472 outages across 268 public firms. Digital platforms account for 18.9% of events, cloud and enterprise software for 15.8%, telecommunications for 13.1%, airlines and travel platforms for 11.6%, payment and financial platforms for 10.9%, logistics and delivery for 9.7%, streaming and entertainment for 8.4%, health-service networks for 5.9%, insurance and account portals for 3.2%, and other recurring service systems for 2.5%. The median outage lasts 5.6 hours, while the 90th percentile lasts 39.4 hours. About 24.3% of events involve multi-region disruption, and 18.7% involve repeat outage at the same firm within the prior six months.

Event timing is based on the earliest public evidence of disruption [6]. If a monitoring feed records failure before firm acknowledgment, the monitoring timestamp is used for outage start. If customer complaints spike before monitoring coverage, the complaint inflection is used. If the first visible signal is a firm status-page update, the update time is used. For daily stock-market tests, the event date is the trading day on which the outage becomes observable. For intraday tests, events are retained only when timestamps are precise within one hour.

The outage severity measure combines duration, affected geography, affected service scope, customer complaint intensity, customer criticality, service-level obligation exposure, and restoration instability [7]. Duration alone is insufficient. A two-hour payment outage during peak transaction hours may be more severe than a twelve-hour outage affecting a low-use entertainment feature [8]. The severity score therefore uses usage intensity and service criticality. Criticality is high for payment access, brokerage trading, health scheduling, travel operations, emergency communication, cloud hosting, and logistics routing.

The outage severity index is:

$$\begin{aligned} OSI_{ie} &= \omega_1 Duration_{ie} + \omega_2 Scope_{ie} \\ &\quad + \omega_3 Complaints_{ie} + \omega_4 Criticality_{ie} \\ &\quad + \omega_5 Instability_{ie}, \\ \sum_{k=1}^{5} \omega_k &= 1. \end{aligned} \tag{1}$$

The weights are estimated from the first principal component of standardized inputs. The component explains 59.8% of the variation. The loading is 0.42 for duration, 0.49 for scope, 0.51 for complaint intensity, 0.47 for criticality, and 0.36 for restoration instability. Equal-weighted versions are used as robustness checks.

The accountability signal is extracted from firm statements during the first seven days after the outage. Statements include status-page updates, press releases, executive posts, customer emails, app notifications, support messages, regulatory notices, and earnings-call remarks when they occur within the response window. The response text is segmented into apology, failure diagnosis, responsibility acceptance, restoration milestone, customer remedy, prevention commitment, recurrence acknowledgment, and uncertainty disclosure. A transformer-based classifier is trained on 76,000 manually labeled statement segments and achieves a holdout macro F1 score of 0.842.

The accountability specificity score is:

$$\begin{aligned} ACC_{ie} &= \lambda_1 Diagnosis_{ie} + \lambda_2 Responsibility_{ie} \\ &\quad + \lambda_3 Timeline_{ie} + \lambda_4 Remedy_{ie} \\ &\quad + \lambda_5 Prevention_{ie} - \lambda_6 Vagueness_{ie}. \end{aligned} \tag{2}$$

The score is standardized by industry and year. Diagnosis indicates whether the firm identifies the failure mechanism at a meaningful level, such as software release error, capacity overload, routing failure, vendor interface failure, authentication service breakdown, database replication issue, scheduling engine failure, or call-center system overload [9]. Responsibility indicates whether the firm acknowledges accountability without shifting blame entirely to customers, weather, unnamed third parties, or unexplained conditions. Timeline indicates whether restoration milestones are provided. Remedy indicates whether the firm specifies customer compensation, fee waiver, service credit, rebooking support, claim process, or operational assistance. Prevention indicates whether the firm describes a concrete action to reduce recurrence.

The response architecture is also coded. A firm can issue a technical update, customer remedy notice, executive apology, regulatory filing, service-credit notice, operational postmortem, or general statement. The study does not assume that one format is superior [10]. Instead, it tests which combinations predict recovery. A technical postmortem can be useful for cloud customers but less effective for stranded travelers. A customer remedy can calm anger but leave investors uncertain about recurrence. The response architecture variables allow these differences to be examined.

Customer outcomes are measured from complaint portals, customer social text, app-store reviews, product forums, direct service-status comments, support-thread volume, search behavior, and app or web usage when available. Complaint volume is normalized by user base or customer count. Sentiment is classified into anger, reliability concern, fairness concern, switching intention, compensation dissatisfaction, and restored confidence [11]. The reliability concern category is especially important because an outage damages reputation when stakeholders infer that failure may recur.

Market outcomes include abnormal stock returns, abnormal volume, bid-ask spread changes, option-implied volatility, put skew, analyst forecast revisions, and analyst report language. Expected stock returns are estimated with a five-factor model plus momentum over trading days minus 260 through minus 31. The primary return window is day minus one through day plus one. Option variables are available for 183 firms with sufficiently liquid options. Analyst data include forecast revisions and report text in the 30 trading days after the outage [12].

Employee text is included because outages create internal strain. Public employee reviews and verified employee discussion are classified into workload pressure, management clarity, technical debt, customer abuse, burnout, and confidence in remediation. Employee signals are not available for all firms, and they are used mainly for supplementary tests. When available, they help determine whether the outage is seen internally as a one-time failure or a symptom of chronic operating stress.

Cao (2022) [13] demonstrates that brand equity can influence firm value through warranty claim rates and abnormal warranty accrual rates, showing that marketing assets can affect valuation through expected cost and reliability channels. This evidence is relevant because outage reputation is also linked to expected remediation cost, service credits, reliability investment, customer support expense, and future service-level penalties [14].

The final panel contains event-level, firm-level, and weekly stakeholder-level observations. For each outage, the dataset records severity, response architecture, accountability specificity, customer sentiment trajectory, complaint decay, app or web activity, market reaction, analyst discussion, recurrence within six months, and recovery time. Recovery time is measured separately for customer sentiment, complaint volume, reliability concern, and implied volatility [15].

3. Measures and Identification Variables

The primary reputation recovery outcome is reliability sentiment recovery. Reliability sentiment is the share of customer text expressing concern that the service is unreliable, fragile, unsafe to depend on, or likely to fail again. Recovery occurs when reliability concern returns to within 10% of its pre-outage baseline and remains there for three consecutive weeks. This measure is more specific than general sentiment because customers can stop being angry while still distrusting reliability [16]. A firm that provides compensation may reduce anger without reducing perceived recurrence risk.

The reliability concern measure is: [17]

$$REL_{it} = \frac{\sum_{m \in M_{it}} q_m P(RelConcern_m)}{\sum_{m \in M_{it}} q_m}. \tag{3}$$

Here, m indexes customer text, q_m weights by credibility and reach, and $P(RelConcern_m)$ is the classifier probability that the message expresses reliability concern. The measure is computed weekly from eight weeks before to twenty-four weeks after the outage. The baseline is the average of weeks minus eight through minus two [18].

The second recovery outcome is complaint half-life. It is the number of weeks required for abnormal complaint volume to fall by 50% from its first-week post-outage level. Complaint half-life is useful because some firms receive immediate spikes that decay quickly, while others face prolonged service anger. The third recovery outcome is churn proxy change. Churn proxies include cancellation searches, account-deletion help-page visits, app uninstall rates, subscriber downgrade language, and service-switching text. These proxies are standardized within industry.

The market uncertainty outcome is implied-volatility persistence. It measures how much of the initial implied-volatility increase remains 20 trading days after the outage. A firm can experience a short-lived volatility increase that disappears after clear restoration, or a persistent increase when investors remain uncertain about recurrence, compensation cost, or system reliability. The persistence measure is:

$$IVP_{ie} = \frac{\Delta IV_{i,e+20}}{\Delta IV_{i,e+2}}, \tag{4}$$

where ΔIV is adjusted for market and industry implied volatility. Values above one indicate that uncertainty increased further after the first response window. Values near zero indicate rapid normalization.

The response credibility variable combines accountability specificity with later consistency. Later consistency is measured by whether the firm's stated restoration timeline and prevention commitment are contradicted by recurrence within 30 days. Because recurrence occurs after the response, it is not used in the main immediate-reaction regressions. It is used in recovery and validation tests. A high-accountability signal followed by repeat outage is less credible ex post and predicts slower long-run reputation recovery.

The response credibility measure is:

$$\begin{aligned} CRD_{ie} = {} & ACC_{ie} - \chi_1 MissedTimeline_{ie} \\ & - \chi_2 RepeatOutage_{ie} - \chi_3 UnreconciledCause_{ie}. \end{aligned} \tag{5}$$

This measure is used for post-event recovery, not for immediate event-window returns. It reflects the idea that accountability becomes more credible when subsequent facts confirm the firm's explanation.

The study distinguishes four response types. Symbolic response is high apology and low operational specificity. Technical response is high diagnosis and low customer remedy. Compensatory response is high remedy and low diagnosis. Integrated accountability response is high diagnosis, responsibility, remedy, timeline, and prevention. These types are not mutually exclusive in raw text, but the classifier assigns each event to the dominant response pattern. Integrated accountability responses account for 27.6% of events, compensatory responses 21.4%, technical responses 18.9%, symbolic responses 20.1%, and minimal or delayed responses 12.0%.

The firm's pre-outage reliability reputation is measured using prior outage history, service-rating sentiment, complaint baseline, analyst discussion of reliability, customer tenure proxies, and service-level claims.

Table 1. Sample Composition by Industry :contentReference[oaicite:0]index=0

Industry	Share (%)	Firms	Events
Digital Platforms	18.9	52	278
Cloud/Enterprise	15.8	41	233
Telecommunications	13.1	33	193
Airlines/Travel	11.6	29	171
Payments/Finance	10.9	27	160
Logistics/Delivery	9.7	24	143

A strong reliability reputation means that stakeholders generally viewed the firm as dependable before the event. This can protect the firm if the outage is seen as isolated, but it can also intensify expectation violation when the outage is severe. The empirical model allows both effects through interactions.

The service criticality variable measures how costly service unavailability is for customers. Payment access, brokerage trading, health scheduling, cloud hosting, travel operations, telecommunications, and logistics routing receive high criticality. Entertainment streaming and nonessential app features receive lower criticality. Criticality moderates both customer reaction and market reaction. A brief outage in a high-criticality service can damage reputation more than a long outage in a low-criticality service.

The recurrence-risk proxy is estimated before the event using prior outage frequency, technical-debt language in employee text, capitalized software intensity, customer complaint baseline, system migration disclosures, and service complexity. This proxy helps distinguish firms with isolated incidents from those with fragile systems. It is included in models predicting recovery and market uncertainty.

The main empirical treatment in machine-learning models is high accountability specificity, defined as an accountability score in the top quartile among firms with comparable outage severity. This severity-adjusted definition matters because severe outages usually require more communication. A firm receives hig accountability treatment only when its response is more specific than expected given the outage's severity, industry, and communication requirements.

The severity-adjusted accountability residual is:

$$\begin{aligned} ACC_{ie} &= \alpha_j + \lambda_t + \delta\, OSI_{ie} \\ &\quad + \Gamma' FirmControls_{ie} + \nu_{ie}, \\ HACC_{ie} &= 1[\widehat{\nu}_{ie} > Q_{75}]. \end{aligned} \tag{6}$$

The high-accountability indicator $HACC$ is used in double machine-learning and causal-forest models. This approach reduces the risk that the estimated effect of accountability merely reflects the fact that more severe outages receive more elaborate responses.

4. Empirical Strategy

The empirical strategy begins with an event-study analysis of market reactions to outages. The first model estimates the effect of outage severity and accountability specificity on abnormal returns. Because communication often occurs after the outage begins, the return model separates immediate outage severity from response information when timestamps allow. For daily tests, response variables are included only if the response occurs within the three-day event window. For recovery tests, response variables are measured over the first seven days [19].

The baseline return model is:

$$\begin{aligned} CAR_{ie} = \alpha_i &+ \lambda_t + \beta_1\, OSI_{ie} \\ &+ \beta_2 ACC_{ie} + \beta_3\, OSI_{ie} ACC_{ie} \\ &+ \Gamma' X_{ie} + \varepsilon_{ie}. \end{aligned} \tag{7}$$

The coefficient on outage severity is expected to be negative. The coefficient on accountability specificity is ambiguous in the immediate return window because a detailed response may reveal greater seriousness. The interaction tests whether accountability softens the penalty of severity. Controls include firm size, leverage, cash, operating margin, prior returns, volatility, analyst coverage, institutional ownership, industry, prior outage history, and media intensity. Standard errors are clustered by firm and event date.

The second model focuses on implied-volatility persistence:

$$\begin{aligned} IVP_{ie} = \alpha_i &+ \lambda_t + \theta_1\, OSI_{ie} \\ &+ \theta_2 ACC_{ie} + \theta_3 Remedy_{ie} \\ &+ \theta_4 Prevention_{ie} + \Theta' X_{ie} + u_{ie}. \end{aligned} \tag{8}$$

This model tests whether accountability reduces uncertainty after controlling for the outage itself. Prevention detail is expected to reduce volatility persistence more than apology language because investors are concerned with recurrence and cost.

The third model estimates customer reputation recovery with an accelerated failure-time specification:

$$\begin{aligned} \log(T_{ie}) = \alpha &+ \rho_1 ACC_{ie} + \rho_2\, OSI_{ie} \\ &+ \rho_3\, Criticality_{ie} + \rho_4\, Compensation_{ie} \\ &+ \rho_5\, RepeatRisk_{ie} + \Omega' Z_{ie} + \xi_{ie}. \end{aligned} \tag{9}$$

Figure 7. Outage market shock.

Table 2. Outage Severity Components and Loadings :contentReference[oaicite:1]index=1

Component	Description	Loading	Weight
Duration	Hours of outage	0.42	0.20
Scope	Geographic/service spread	0.49	0.22
Complaints	Customer intensity	0.51	0.23
Criticality	Service importance	0.47	0.20
Instability	Restoration volatility	0.36	0.15

Here, T_{ie} is the number of weeks until reliability sentiment recovery. A negative coefficient means faster recovery. The model is estimated with log-logistic and Weibull distributions; the reported results are similar [20]. Censored events are those that do not recover within 24 weeks.

The fourth model uses matched synthetic controls for stakeholder outcomes. Each outage firm is matched to a synthetic combination of non-outage firms in the same industry with similar pre-event customer sentiment, app usage, complaint volume, market volatility, and size. The synthetic design estimates the abnormal path of customer sentiment and complaint volume relative to a counterfactual without outage [21]. This design is used for 823 events with sufficient pre-event stakeholder data [22].

The synthetic sentiment gap is:

$$Gap_{i,e+h} = Y_{i,e+h} - \sum_{k\in C_i} w_k Y_{k,e+h},$$
$$\sum_{k\in C_i} w_k = 1, \quad w_k \geq 0. \tag{10}$$

The weights are chosen to match pre-outage stakeholder and firm trajectories. The gap is then compared across response types [23]. This approach does not prove random assignment, but it provides a transparent recovery benchmark.

The fifth model uses double machine learning to estimate the association between high accountability and outcomes after flexible control. The outcomes are reliability recovery within six weeks, complaint half-life, implied-volatility persistence, abnormal returns, and churn proxy change. The treatment is severity-adjusted high accountability. Nuisance models predict treatment and outcome using outage severity, industry, firm characteristics, prior reliability reputation, recurrence risk, media attention, and customer baseline.

The orthogonal estimator is: [24]

$$\begin{aligned}\widetilde{Y}_{ie} &= Y_{ie} - \widehat{m}(W_{ie}),\\ \widetilde{D}_{ie} &= D_{ie} - \widehat{g}(W_{ie}),\\ \widehat{\tau} &= \frac{\sum_{i,e}\widetilde{D}_{ie}\widetilde{Y}_{ie}}{\sum_{i,e}\widetilde{D}_{ie}^2}.\end{aligned} \tag{11}$$

Cross-fitting is conducted by firm. The covariate vector includes all variables observed before the response window, including severity, industry, pre-event customer sentiment, prior outage frequency, operating margin, service criticality, media environment, and baseline complaint rate. This timing restriction avoids conditioning on post-response recovery variables.

The final model uses causal forests to estimate heterogeneity. The treatment is high accountability. Outcomes are six-week reliability recovery and implied-volatility persistence. Moderators include outage severity, service criticality, prior reliability reputation, repeat-

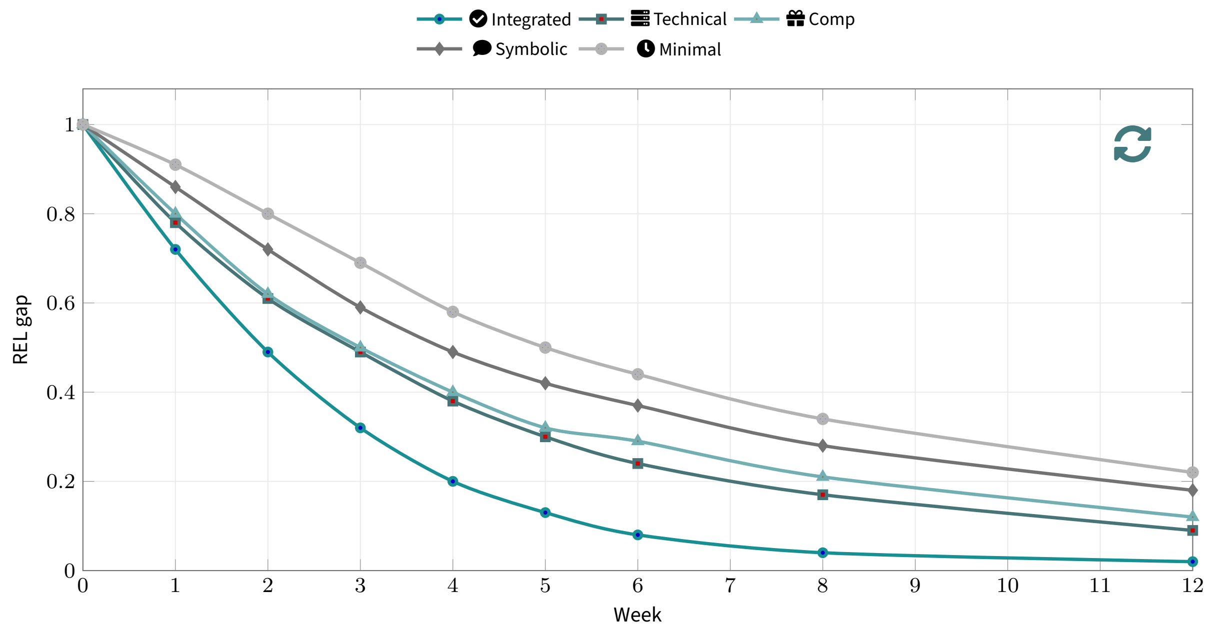

Figure 8. Reliability recovery.

Table 3. Accountability Signal Components :contentReference[oaicite:2]index=2

Component	Meaning	Effect on Recovery	Role
Diagnosis	Identifies failure cause	High	Core
Responsibility	Accepts accountability	Medium	Trust
Timeline	Restoration milestones	High	Clarity
Remedy	Customer compensation	Medium	Fairness
Prevention	Future safeguards	Very High	Credibility

risk proxy, customer switching cost, industry, employee technical-debt language, complaint intensity, and response format. The heterogeneous effect is:

$$Y_{ie} = m(W_{ie}) + \tau(W_{ie})D_{ie} + \epsilon_{ie},$$
$$\tau(W_{ie}) = E[Y(1) - Y(0) \mid W_{ie}]. \quad (12)$$

The causal forest is interpreted as an adjusted heterogeneity model, not as randomized evidence. Firms choose response strategies, and response quality may reflect unobserved competence. The analysis addresses this through severity adjustment, fixed effects, matched comparisons, and recurrence validation, but the interpretation remains disciplined.

5. Results

The average outage produces a negative market reaction. The mean three-day abnormal return is minus 0.76%, with a median of minus 0.41%. High-severity outages generate a mean abnormal return of minus 1.64%, while low-severity outages generate minus 0.22%. The negative reaction is strongest in payments, brokerage, health-service networks, telecommunications, airlines, cloud services, and logistics. Streaming and entertainment outages produce smaller return effects unless they occur during major releases or subscriber reporting periods.

Outage severity predicts returns. A one-standard-deviation increase in the outage severity index is associated with a 0.52 percentage point lower three-day abnormal return after firm fixed effects, calendar effects, industry controls, prior outage history, and media intensity are included. Service criticality strengthens the reaction. A severe outage in a high-criticality service produces nearly double the market penalty of a similarly long outage in a low-criticality service. Duration alone is less predictive than the full severity index.

Accountability specificity moderates the penalty of severity [25]. In immediate event-window models, the standalone coefficient on accountability is small because detailed responses often follow severe outages. The interaction between severity and accountability is positive. High-accountability responses reduce the return penalty of severe outages by about 0.31 percentage points. The moderation is stronger when responses occur within the first trading day. Delayed high-specificity responses still help reputation recovery but have weaker immediate return effects.

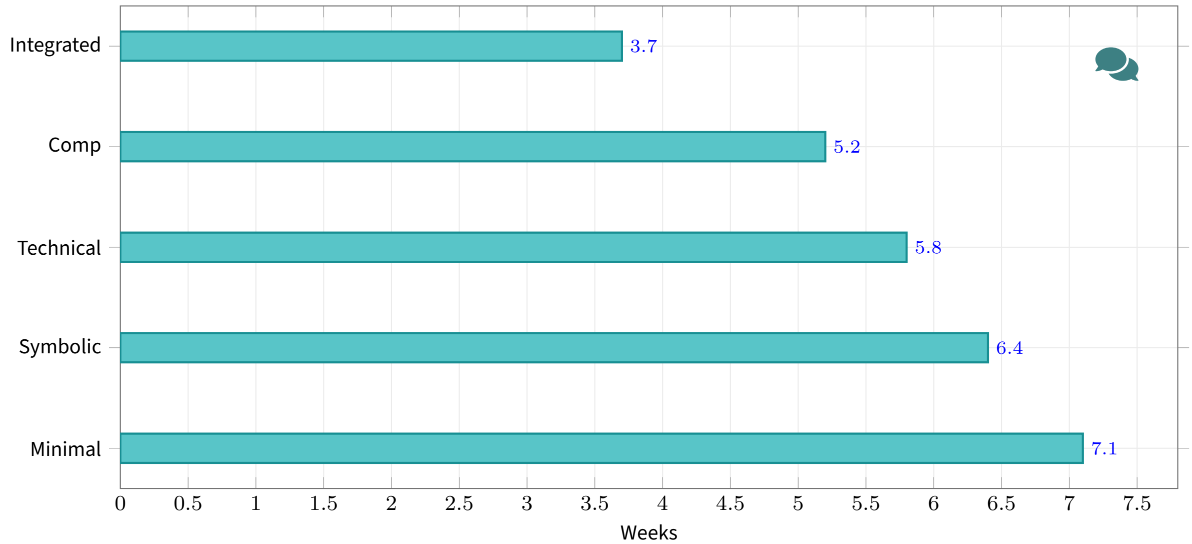

Figure 9. Complaint half-life.

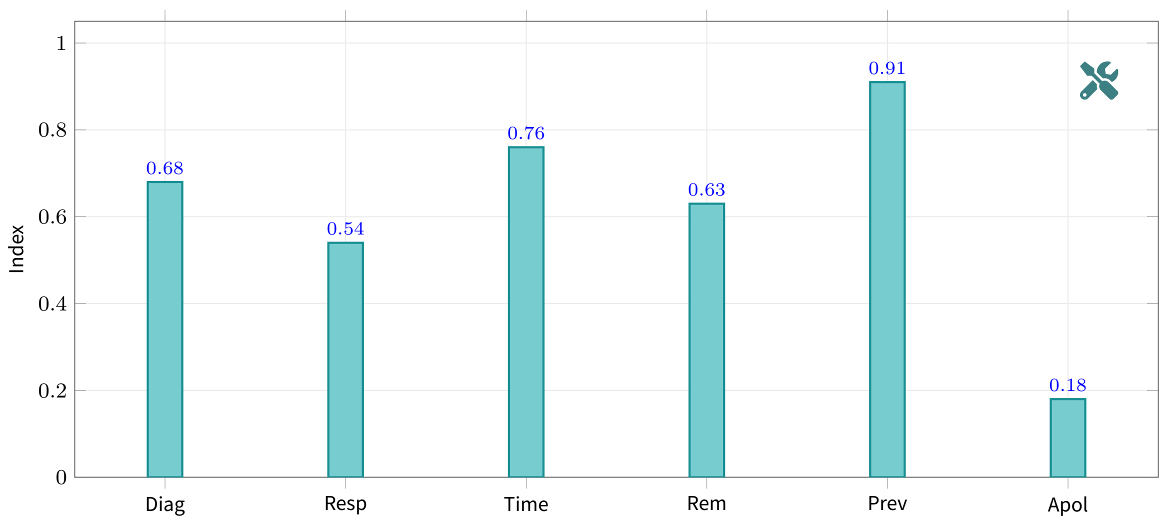

Figure 10. Component strength.

Option markets show a clearer accountability effect. Thirty-day implied volatility rises by 3.08% after the average outage and by 6.42% after high-severity outages. High-accountability responses reduce implied-volatility persistence by 2.18 percentage points after controls. Prevention commitments and restoration milestones are the strongest components. Apology language does not significantly reduce volatility persistence once diagnosis, timeline, and prevention are included. Investors appear to care more about recurrence uncertainty than about regret.

Customer complaints surge immediately. Complaint volume rises 64.2% in the first week after the average outage. High-severity outages produce a 128.5% increase. The complaint half-life is 3.7 weeks for integrated accountability responses, 5.2 weeks for compensatory responses, 5.8 weeks for technical responses, 6.4 weeks for symbolic responses, and 7.1 weeks for minimal or delayed responses. The difference remains after severity adjustment. Compensation reduces complaint volume quickly, but the effect fades if reliability concern remains high.

Reliability sentiment recovers more slowly than general sentiment. General negative sentiment returns to baseline in a median of 5.1 weeks. Reliability concern recovers in a median of 6.8 weeks. For high-severity outages, reliability recovery takes 9.6 weeks. Integrated accountability responses reduce median reliability recovery time by 39.4%. Symbolic responses reduce recovery time by only 8.7%, and the estimate is not statistically stable. Technical responses reduce analyst uncertainty but have weaker customer recovery unless paired with customer remedy.

The accelerated failure-time model confirms these

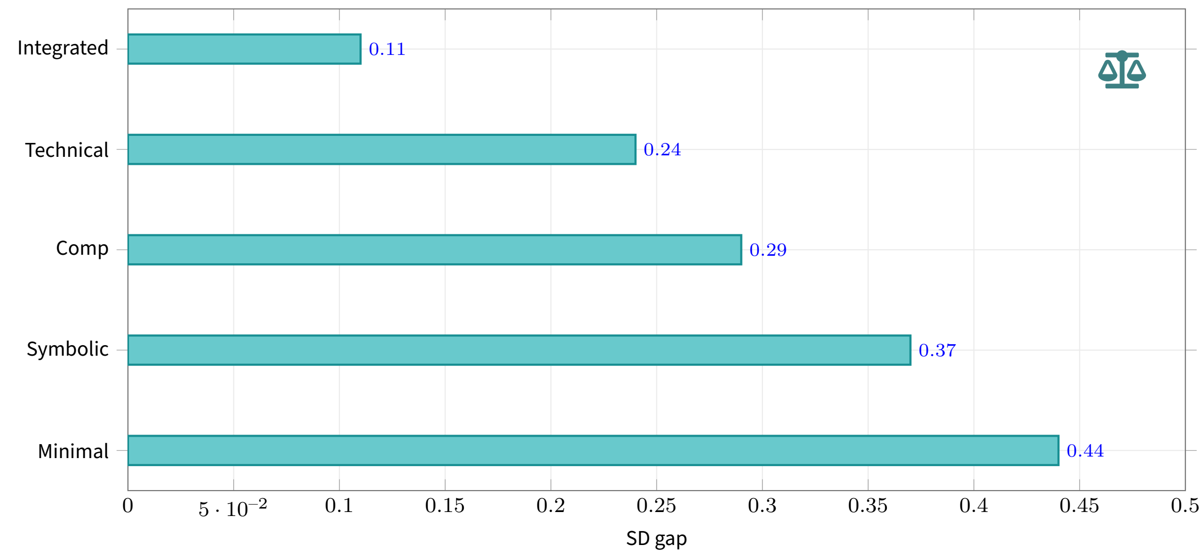

Figure 11. Synthetic baseline gap.

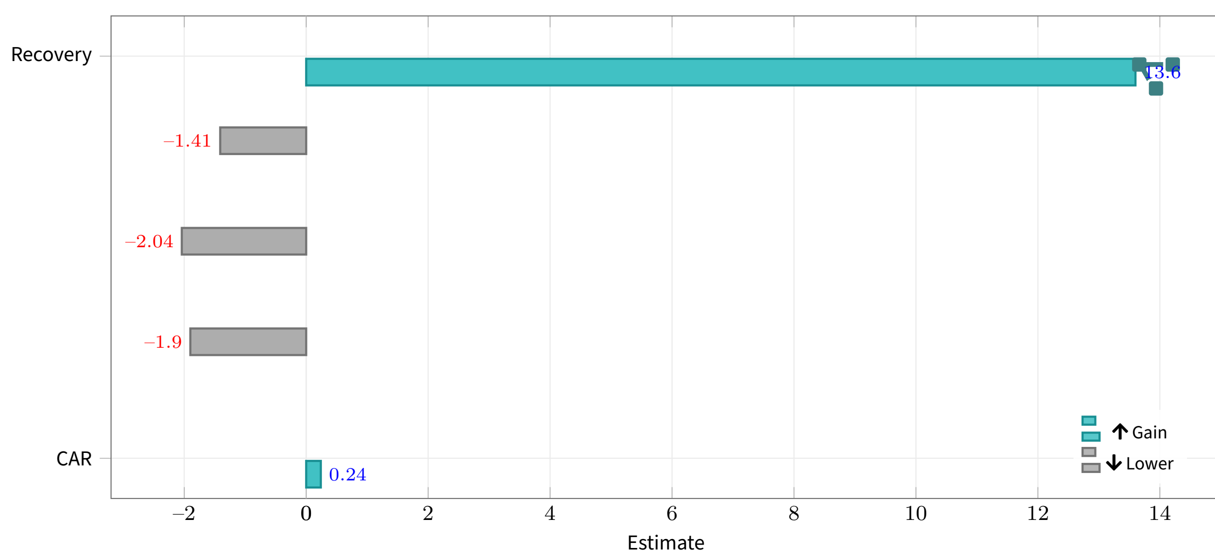

Figure 12. DML effect estimates.

differences. The coefficient on accountability specificity is minus 0.502, implying a 39.4% reduction in median reliability recovery time. Outage severity increases recovery time, service criticality increases recovery time, and repeat-risk proxy increases recovery time. Compensation has a short-run complaint effect but a weaker reliability recovery effect. Prevention detail has the largest association with reliability recovery among response components. Customers appear to require evidence that recurrence risk is addressed [26].

The response architecture results are consistent with the component analysis. Integrated accountability responses generate the fastest recovery. Compensatory responses are effective for low-criticality services and billing-related outages but less effective for access failures in critical services. Technical responses work well for enterprise cloud customers and analysts but less well for mass-market customers unless translated into customer-facing implications. Symbolic responses generate media coverage but little measurable recovery unless the outage is low severity. Minimal responses produce the slowest recovery and largest complaint persistence.

Executive apologies receive attention but are not sufficient. Events with chief executive or founder apologies have higher media intensity and larger initial complaint volume, reflecting the severity of events that trigger executive involvement. After severity adjustment, executive apology predicts faster recovery only when the message includes diagnosis, remedy, timeline, and prevention. Executive apology without operational detail has no significant effect on reliability recovery and

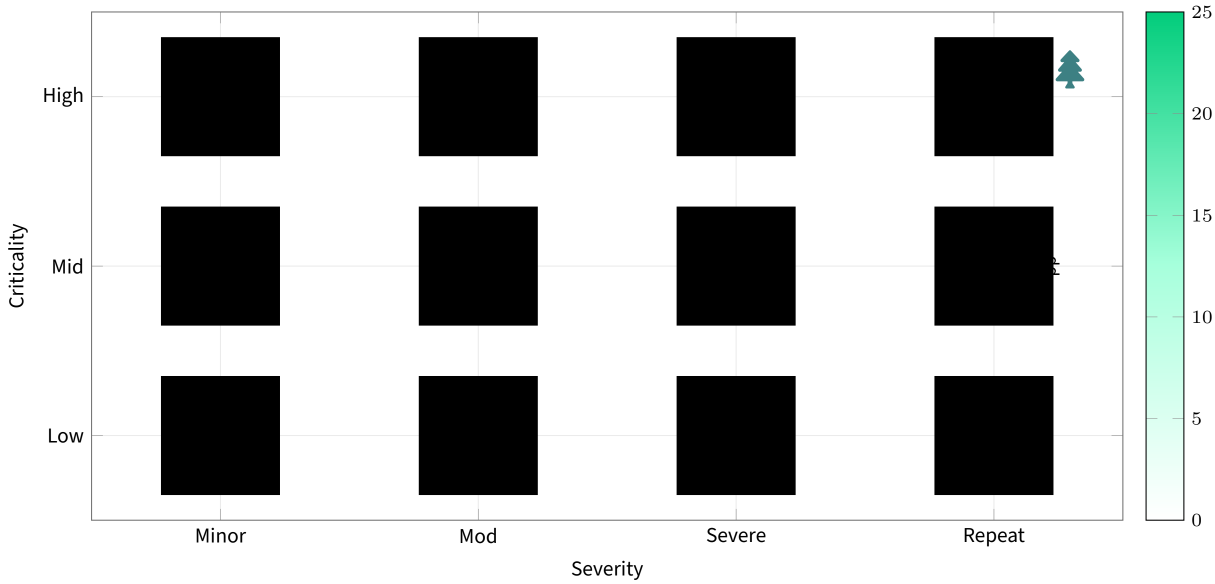

Figure 13. Forest heterogeneity.

Table 4. Market Reaction Summary :contentReference[oaicite:3]index=3

Metric	Mean	High Severity	Low Severity
CAR (-1,+1)	-0.76%	-1.64%	-0.22%
IV Increase	3.08%	6.42%	1.12%
Volume Change	+18.4%	+32.7%	+6.5%
Spread Change	+9.2%	+15.8%	+3.1%

is associated with higher implied-volatility persistence in high-criticality services. The market may interpret a high-profile apology without detail as a sign that the firm lacks a clear technical resolution.

Cao and Xu (2022) [27] find that brand transaction outcomes can depend on managerial valuation and decision discipline, which is relevant here because outage response is also a managerial signal under uncertainty. In this setting, investors appear to reward response discipline when managers convert a visible failure into a verifiable recovery plan rather than relying on symbolic communication.

Analyst reports respond to outage content and response content. After high-severity outages, analyst reports increase mentions of reliability risk by 28.6%, service credit cost by 17.4%, churn risk by 14.9%, and infrastructure investment by 22.1%. High-accountability responses reduce follow-up analyst concern about recurrence but increase discussion of remediation cost. This tradeoff is important. A credible prevention plan may reassure investors about reliability while revealing future spending. The net abnormal return effect depends on whether recurrence concern or cost concern dominates.

Customer churn proxies increase after severe outages. Cancellation searches, account-deletion visits, uninstall rates, and switching-language frequency rise by 6.8% over four weeks after high-severity outages. Integrated accountability responses reduce this increase by 2.3 percentage points. Compensation-only responses reduce cancellation searches in the first two weeks but not uninstall rates or switching language in weeks three through six. This suggests that compensation can delay exit but does not fully restore confidence.

The synthetic-control results validate the recovery patterns. For high-severity outages with minimal responses, reliability sentiment remains 0.44 standard deviations above synthetic baseline at week six. For symbolic responses, the gap is 0.37. For compensatory responses, the gap is 0.29. For technical responses, the gap is 0.24. For integrated accountability responses, the gap is 0.11. The ranking remains similar at week twelve, though all gaps shrink. This confirms that response architecture predicts the speed of reputation normalization.

Double machine-learning estimates show that severity-adjusted high accountability increases the probability of reliability recovery within six weeks by 13.6 percentage points. It reduces complaint half-life by 1.41 weeks, reduces implied-volatility persistence by 2.04 percentage points, and reduces churn proxy growth by 1.9 percentage points [28]. The effect on immediate abnormal return is positive but smaller, at 0.24 percentage points, and statistically weaker. The accountability

Table 5. Customer Response Metrics :contentReference[oaicite:4]index=4

Metric	Average	High Severity	Low Severity
Complaint Spike	64.2%	128.5%	21.3%
Sentiment Drop	41.7%	73.2%	15.6%
Recovery Time (weeks)	6.8	9.6	4.2
Churn Proxy Increase	6.8%	11.4%	2.1%

Table 6. Response Type and Recovery Speed :contentReference[oaicite:5]index=5

Response Type	Share (%)	Half-life (weeks)	Recovery Speed
Integrated	27.6	3.7	Fastest
Compensatory	21.4	5.2	Moderate
Technical	18.9	5.8	Moderate
Symbolic	20.1	6.4	Slow
Minimal/Delayed	12.0	7.1	Slowest

signal has clearer effects on recovery and uncertainty than on immediate price.

The causal forest identifies where accountability matters most. The largest recovery effects occur for high-criticality services, high pre-outage reliability reputation, moderate outage severity, high switching costs, and low recurrence within 30 days. Under these conditions, high accountability increases six-week reliability recovery probability by 22.4 percentage points. The smallest effects occur for very low-criticality services, extremely severe outages, and firms with repeat outages shortly after the response. For extreme outages, communication helps but cannot overcome operational evidence of fragility.

Prior reliability reputation has a nonlinear effect. Firms with strong reliability reputation experience less initial customer anger for moderate outages, but more severe expectation violation for major outages. High-accountability responses are especially valuable for these firms because stakeholders are willing to interpret the outage as unusual if management provides a credible explanation. Firms with weak prior reliability reputation receive smaller gains from accountability because stakeholders are less willing to believe that the failure is isolated. Reputation protects only when the response supports the interpretation of exceptionality.

Recurrence damages credibility. Among firms issuing high-accountability responses, those experiencing a repeat outage within 30 days have much slower recovery. Their reliability concern remains 0.31 standard deviations above baseline at week twelve, compared with 0.09 for firms without recurrence. The market also penalizes recurrence more strongly when the firm had made a prevention commitment. This result confirms that accountability signals are not cheap talk when later performance can validate or falsify them.

Employee text provides supplementary evidence [29]. After severe outages, employee discussion of workload pressure and customer abuse rises sharply. Firms with integrated accountability responses show faster improvement in employee confidence in remediation. Firms with symbolic responses show persistent employee comments about unclear priorities and technical debt. These patterns are not central to the financial tests, but they suggest that internal stakeholders also react to the credibility of outage response.

6. Additional Tests and Robustness

The first robustness test changes the outage severity definition. A stricter definition requiring both high duration and high complaint intensity reduces the sample to 914 events. The market penalty becomes larger, with a mean abnormal return of minus 1.08%, and the accountability recovery effect remains strong. A broader definition including minor outages increases the sample to 2,311 events and reduces the average penalty to minus 0.29%. The core relationship between accountability and recovery remains present but weaker. This pattern supports the idea that accountability matters most when the outage is visible enough to threaten reputation.

The second test uses alternative event windows. The two-day abnormal return is minus 0.61%, the three-day return is minus 0.76%, and the five-day return is minus 0.89%. The eleven-day return is minus 1.02% but includes more confounding information. Accountability moderation is strongest in the five-day window because some firms respond after the first day. Intraday tests show that when firms issue high-specificity statements within three hours of outage recognition, bid-ask spreads normalize faster and abnormal volume decays sooner.

The third test uses alternative expected-return models. Market-adjusted returns, industry-adjusted returns, a four-factor model, and a five-factor model with momentum all produce negative outage severity coefficients and positive severity-accountability interactions.

Table 7. Regression Outcome Summary :contentReference[oaicite:6]index=6

Variable	CAR Effect	IV Persistence	Recovery Time
Severity (OSI)	Negative	Positive	Slower
Accountability	Mild Positive	Negative	Faster
Prevention	Neutral	Strong Negative	Fastest
Compensation	Short-term +	Neutral	Limited
Apology	Insignificant	Insignificant	Weak

Table 8. Machine Learning Performance :contentReference[oaicite:7]index=7

Model	Baseline AUC	+Outage Vars	+Accountability
Recovery Prediction	0.641	0.734	0.812
IV Persistence	0.659	0.746	0.789
Return Prediction	0.702	0.755	0.776
Churn Prediction	0.618	0.701	0.768

The magnitude of the average outage penalty ranges from minus 0.68% to minus 0.82%. The preferred estimate is in the middle of the range. The findings are not dependent on a single asset-pricing specification.

The fourth test excludes outages occurring near earnings announcements, investor days, product launches, or regulatory filings. The restricted sample contains 1,038 events. The average abnormal return is minus 0.69%, and accountability specificity still predicts faster reliability recovery. This reduces concern that the results are driven by overlapping corporate news.

The fifth test separates first outages from repeated outages. First outages generate an average abnormal return of minus 0.54%. Repeated outages within six months generate minus 1.23%. Accountability responses are less effective after repeated outages, especially when the firm previously made prevention commitments. Stakeholders appear to update from isolated failure to systematic fragility when outages recur. Repetition shifts the meaning of communication from explanation to credibility test [30].

The sixth test examines compensation size [31]. Larger compensation packages reduce complaint volume but can increase analyst cost concern. In payments, airlines, travel, and cloud services, service credits and refunds are associated with faster complaint decay but lower near-term margin forecasts. The market reaction is most favorable when compensation is paired with prevention detail. Compensation without prevention is interpreted as cost without reliability improvement.

The seventh test studies technical detail. Excessive technical detail without translation reduces analyst uncertainty but has little customer benefit [32]. Customer-facing reliability sentiment improves when technical diagnosis is paired with plain-language explanation of what users should expect next. The best-performing responses include enough detail to be credible but enough customer relevance to be understandable. This result suggests that communication must match stakeholder information needs.

The eighth test investigates response channel. Status-page updates are effective for enterprise software and cloud customers but less visible to mass-market customers. App notifications reduce complaint volume faster for consumer platforms. Executive social posts increase attention but are risky when they omit remedy detail. Customer emails improve recovery when they contain individualized remedy information. Press releases are less effective for immediate customer recovery but useful for investor interpretation. Channel matters because stakeholders receive information differently.

The ninth test examines media framing. Outages framed as technical accidents produce smaller reputation losses than outages framed as managerial underinvestment, outdated systems, or ignored warnings. Accountability specificity reduces negative reaction most when media framing emphasizes uncertainty. It reduces reaction less when media framing emphasizes long-term underinvestment. The firm's response is less persuasive when external narratives already define the outage as predictable.

The tenth test analyzes analyst questions in subsequent earnings calls. Outage firms receive more questions about reliability investment, redundancy, service credits, customer churn, and system modernization [33]. High-accountability responses reduce the number of recurrence questions but increase questions about capital expenditure [34]. This confirms the tradeoff observed in analyst reports. Clear prevention reduces uncertainty but makes spending implications salient.

The eleventh test examines customer switching costs. High switching costs reduce actual churn proxies but not reliability concern. Customers may stay because switching is difficult while continuing to express distrust. Investors distinguish between these outcomes. Stock penalties are smaller when switching costs are high, but implied volatility still rises if reliability con-

Table 9. Heterogeneity of Accountability Effects :contentReference[oaicite:8]index=8

Condition	Effect Size	Recovery Gain	Notes
High Criticality	Large	+22.4 pp	Strong impact
Moderate Severity	Largest	+24.7 pp	Optimal zone
Low Criticality	Small	+6.2 pp	Limited need
Repeat Outages	Negative	-12.1 pp	Credibility loss
High Switching Cost	Medium	+15.3 pp	Retention aided

cern persists. This indicates that switching costs protect revenue more than reputation.

The twelfth test uses actual churn or subscriber data for a smaller subset of firms. In this subset, high-severity outages increase next-quarter churn by 0.38 percentage points [35]. High-accountability responses reduce the increase by 0.14 percentage points. The estimate is modest but consistent with proxy results. Effects are largest in subscription software, streaming, and consumer finance apps. Travel and airline churn is harder to measure because purchase frequency is irregular.

The thirteenth test estimates effects by customer type. Enterprise-facing outages generate larger analyst and option reactions than consumer sentiment reactions. Consumer-facing outages generate larger social sentiment and complaint reactions. Hybrid firms experience both. Integrated accountability works across both groups but through different components. Enterprise customers respond more to diagnosis and prevention. Consumers respond more to remedy and timeline. Investors respond to prevention and recurrence evidence.

The fourteenth test removes outages caused primarily by external infrastructure providers. These events may be less attributable to the focal firm. The accountability effect remains but is smaller. Firms that clearly explain third-party dependence and customer protection plans recover faster than firms that simply blame vendors. Stakeholders appear to accept external causes only when the firm explains why customer impact will be reduced in the future.

The fifteenth test examines regulatory involvement. Outages involving regulatory notification or public agency attention produce larger market penalties and slower reputation recovery. Accountability specificity helps, but prevention commitments must be verifiable. In regulated services, vague assurances are especially ineffective. Regulators and analysts both look for operational controls, redundancy, and reporting obligations. The reputational burden is higher when service failure has public-interest implications.

The sixteenth test evaluates whether firms overpromise restoration. Some firms provide specific timelines and then miss them. Missed timelines sharply reduce credibility and slow recovery. A missed restoration milestone increases reliability recovery time by 31.2%. Conservative but accurate timelines perform better than optimistic timelines that later fail. This result highlights that specificity is valuable only when credible. Precision without delivery becomes reputationally costly.

The seventeenth test analyzes postmortem reports. Firms issuing detailed postmortems within 14 days experience lower implied-volatility persistence and fewer analyst recurrence questions. Customer sentiment improves only when the postmortem includes customer remedy and plain-language prevention. Technical postmortems alone mainly benefit investor and enterprise audiences. Public postmortems are most effective after cloud, software, payment, and logistics outages.

The eighteenth test examines capital investment after outages. Severe outages are followed by higher capitalized software expenditure, infrastructure spending, or technology operating expense in the next two quarters. The increase averages 2.1% of quarterly technology-related expense. Firms with high-accountability responses show larger disclosed reliability investments. This supports the interpretation that high-accountability responses often accompany real remediation rather than communication alone. However, spending increases also reduce short-term margin forecasts.

The nineteenth test uses placebo events based on routine scheduled maintenance announcements. Scheduled maintenance does not generate negative abnormal returns or complaint spikes when announced clearly [36]. Unplanned outages do. This distinction supports the interpretation that the market reacts to unexpected reliability failure, not to service unavailability in general. Planned downtime can be reputation-neutral when expectations are managed.

The twentieth test uses non-outage negative news matched by media intensity. Generic negative media attention produces weaker reliability sentiment effects and smaller implied-volatility responses. This confirms that outage events have distinctive informational content. Stakeholders react specifically to reliability failure and recurrence uncertainty, not merely to negative publicity.

7. Machine Learning Validation

The predictive analysis examines whether accountability and outage variables improve out-of-sample prediction of reputation recovery and market uncertainty. The baseline model uses firm financials, industry, prior returns, volatility, analyst coverage, and media intensity. The outage model adds severity, duration, scope, criticality, and complaint intensity. The accountability model adds diagnosis, responsibility, timeline, remedy, prevention, vagueness, response channel, and response delay.

For predicting reliability recovery within six weeks, the baseline model achieves an out-of-sample area under the curve of 0.641. Adding outage variables increases it to 0.734. Adding accountability variables increases it to 0.812. The improvement is substantial because recovery depends not only on outage severity but also on how the firm reduces uncertainty afterward. The strongest predictors are prevention detail, restoration timeline accuracy, service criticality, complaint intensity, prior outage frequency, remedy specificity, and recurrence within 30 days when predicting later recovery.

For predicting implied-volatility persistence, the baseline model achieves an area under the curve of 0.659. Outage variables raise it to 0.746. Accountability variables raise it to 0.789. The most important predictors are outage criticality, recurrence-risk proxy, prevention detail, technical diagnosis, analyst coverage, and response delay [37]. Apology language has low predictive importance after operational content is included. This confirms the regression evidence that markets care more about reliability uncertainty than symbolic regret.

For predicting abnormal returns below minus 2%, the full model achieves an area under the curve of 0.776. The strongest predictors are outage severity, service criticality, prior reliability reputation, repeat outage history, customer complaint intensity, and whether response occurs before market close. Accountability specificity helps prediction but less than for recovery outcomes because the immediate return window reflects the outage itself more strongly than the response. This difference clarifies why accountability has stronger effects on recovery and volatility persistence than on immediate abnormal return.

Causal-forest heterogeneity confirms that accountability is most valuable in intermediate conditions. When outages are very minor, reputation recovers quickly regardless of response [38]. When outages are catastrophic or repeated, communication cannot fully offset operational evidence. The largest accountability effect occurs for moderate-to-severe outages in high-criticality services where customers require reassurance but the failure remains plausibly fixable. In this region, high accountability increases six-week reliability recovery probability by 24.7 percentage points.

Variable interaction plots show that prevention detail and remedy specificity are complements. Prevention detail without remedy helps analysts but leaves customers dissatisfied [39]. Remedy without prevention reduces anger but leaves reliability concern. The combination produces the largest recovery effect. Diagnosis and timeline are also complements. Diagnosis explains what failed; timeline explains how restoration is progressing. Without timeline, diagnosis can seem detached from customer experience. Without diagnosis, timeline can seem unsupported.

The model identifies false positives and false negatives. False positives are events predicted to recover quickly but remaining impaired. Manual inspection shows that many involve repeat outages after strong initial communication, missed restoration timelines, or external investigations [40]. False negatives are events predicted to recover slowly but recovering quickly. These often involve loyal customer bases, high switching costs, or unusually generous remedies. The errors reveal that later operational validation remains crucial.

The predictive model also estimates the value of response timing. The marginal benefit of high accountability is largest when the response begins within 12 hours. A late but specific response still improves recovery relative to late vagueness, but the effect is smaller. Early vague responses are less useful than moderately delayed specific responses. Timing is therefore secondary to content but still important when the outage is highly visible.

The machine-learning results support the main conclusion. Accountability signals are not reducible to apology, compensation, or technical disclosure alone. The response must jointly reduce uncertainty about cause, customer harm, restoration, and recurrence. Predictive performance improves most when these elements are modeled together. This is consistent with the idea that reputation recovery after outage depends on credible accountability architecture.

8. Discussion

The results indicate that large service outages are corporate signals because they reveal the condition of reliability systems that stakeholders cannot otherwise inspect. The outage itself signals possible fragility [41]. The response signals managerial accountability and remediation capacity. Stakeholders then observe whether later restoration and recurrence validate the signal. Reputation recovery depends on the entire sequence. A strong response cannot fully repair a weak system, but a weak response can prolong the damage from a repairable outage.

The evidence clarifies why apology alone is often insufficient. An apology addresses relational harm, but outages create competence uncertainty. Customers want to know whether the service will fail again. Investors want to know whether remediation costs, churn, service credits, or infrastructure spending will rise. Employees want to know whether leadership understands the operational problem. A useful accountability signal must therefore connect regret to diagnosis, remedy, and prevention. Without this connection, regret remains symbolically visible but informationally thin.

The findings also show that compensation is not a complete substitute for explanation. Compensation reduces immediate complaint volume because it addresses inconvenience and fairness. Yet reliability sentiment recovers only when customers also receive evidence that recurrence risk is lower. In high-criticality services, customers value future dependability more than one-time credits. The strongest recovery occurs when compensation is paired with plain-language explanation and credible prevention [42].

For managers, the implication is that outage response should be designed before outages occur. Firms should maintain response protocols that identify who communicates, what diagnostic information can be shared, how customer remedies are triggered, how restoration milestones are updated, and how prevention commitments are verified. A firm improvising during a severe outage is more likely to issue vague or inconsistent statements [43]. Prepared accountability architecture can reduce both reputational and market uncertainty.

For investor relations, the results indicate that technical reliability is financially material even when outages are short. Investors react to service criticality, repeat history, prevention credibility, and recurrence. A firm should therefore explain not only that service has been restored but also whether the outage changes capital expenditure, operating expense, service-level credits, customer churn, or system redundancy plans. Analysts respond favorably to clarity about recurrence risk, even when the clarity reveals some remediation cost.

For customer communication, the findings indicate that channel choice matters. Status pages are not enough for mass-market customers. Press releases are not enough for users waiting for access. Executive social posts are not enough for enterprise customers requiring technical detail. Effective communication must match the stakeholder's information need. The same outage may require a technical postmortem for analysts and enterprise users, a remedy notice for customers, and an internal message for employees.

For reputation theory, the paper shows that corporate reputation is maintained through operational evidence as well as symbolic identity [44]. A firm can have a strong brand, but if stakeholders depend on continuous access, reliability is central to reputation. Outages test that reputation under conditions of observable failure. The response then determines whether stakeholders interpret the failure as an isolated incident, a manageable capacity problem, a governance weakness, or evidence of chronic fragility.

For researchers, the paper highlights the value of outage telemetry. Many reputation studies rely on news events or stakeholder text alone. Outage telemetry provides an independent measure of operational failure. This allows response quality to be studied after controlling for the underlying severity of the event. Future work can apply similar designs to product downtime, claims-processing failures, delivery network disruption, and account-access failures in regulated services.

The study has limitations. Outage severity is measured from public and semi-public data, which may not capture all affected customers. App and web activity data are incomplete for some firms. Churn proxies are indirect except in a smaller subset. Firms choose response strategies, so high accountability may partly reflect stronger unobserved operating capability. The analysis uses severity adjustment, fixed effects, matched controls, and machine learning, but it cannot fully separate communication from underlying competence.

The results should therefore be interpreted as evidence that high-specificity accountability responses are associated with faster recovery and lower uncertainty, especially when subsequent operations do not contradict the response. This interpretation is practical. Stakeholders care about the combination of communication and competence [45]. A response that accurately describes real remediation is useful. A response that promises prevention and is followed by repeat outage is damaging [46]. Accountability is valuable because it can be verified.

Future research could use internal incident logs, service-level agreement records, customer-level churn, call-center transcripts, and infrastructure investment data to estimate mechanisms more directly. Cross-country studies could examine whether outage accountability matters more in regulated markets or in countries with stronger consumer protection. Experimental work could test how customers interpret different combinations of apology, remedy, diagnosis, and prevention. Longitudinal studies could examine whether repeated high-accountability responses build resilience or whether repeated outages eventually overwhelm communication credibility.

9. Conclusion

This paper examined how firms use accountability signals after large-scale service outages and how those signals relate to reputation recovery and market reactions. The evidence shows that outages produce negative abnormal returns, higher implied volatility, customer complaint spikes, and persistent reliability concerns. These effects are strongest in high-criticality services where customers depend on continuous access, transaction completion, or operational coordination.

The main finding is that accountability specificity accelerates reputation recovery. Responses that identify the failure mechanism, accept responsibility, provide restoration milestones, specify customer remedies, and commit to prevention shorten reliability sentiment recovery and reduce complaint persistence. Apology alone is not enough. Compensation alone is not enough. Technical explanation alone is not enough [47]. The strongest recovery occurs when these elements are integrated into a coherent accountability signal.

The market evidence shows that investors respond to both outage severity and recurrence uncertainty [48]. Severe outages reduce abnormal returns and raise implied volatility. High-accountability responses reduce volatility persistence more clearly than they improve immediate abnormal returns. This indicates that accountability works primarily by reducing uncertainty about future failures, remediation costs, and customer loss rather than by erasing the initial operating shock.

The stakeholder evidence shows that different audiences require different information. Customers react to remedy, timeline, and plain-language prevention [49]. Analysts react to diagnosis, recurrence risk, infrastructure investment, and service-credit cost. Employees react to management clarity and technical-debt acknowledgment. A credible outage response must therefore address multiple information needs without becoming fragmented or evasive.

The study also shows that accountability signals are validated by subsequent operations. Firms that issue specific prevention commitments and then experience repeat outages recover slowly and face renewed market penalties. Specificity is beneficial only when it is credible. The best response is not the most elaborate statement but the most accurate, actionable, and verifiable explanation of restoration and prevention.

The findings suggest that operational reliability should be treated as a central component of reputation in service-intensive and digital firms. Stakeholders judge firms not only by whether failures occur, but also by whether managers explain them clearly, remedy customer harm, and reduce recurrence risk. As firms become more dependent on continuous digital and service infrastructure, outage accountability will remain a consequential corporate signal. The present evidence indicates that large service outages are not merely technical interruptions. They are public moments in which corporate reputation is revised through the interaction of failure severity, stakeholder harm, managerial accountability, and verified recovery.

References

[1] M. Cano-Rubio, R. Lombardi, G. Fuentes-Lombardo, and P. Núñez-Cacho, "Familiness, business strategy and stakeholder engagement: The internationalisation of spanish olive oil mills," Business Strategy and the Environment, vol. 30, pp. 4258–4280, 8 2021.

[2] F. Murshed, Z. Cao, K. Savitskie, and S. Sen, "Ethical csr, organizational identification, and job satisfaction: Mediated moderated role of interactional justice," Social Justice Research, vol. 36, no. 1, pp. 75–102, 2023.

[3] C. Brown, "Environmental management in the film and television production industry," 5 2021.

[4] J. Connell and J. Burgess, Developments in the Call Centre Industry. Routledge, 9 2006.

[5] Z. Cao and R. Yan, "Health creates wealth? the use of nutrition claims and firm financial performance," Journal of Public Policy & Marketing, vol. 35, no. 1, pp. 58–75, 2016.

[6] M. Fernando and G. Moore, "Macintyrean virtue ethics in business: A cross-cultural comparison," Journal of Business Ethics, vol. 132, pp. 185–202, 8 2014.

[7] A. Wahlman, Exploring Change toward Sustainability in Established Business Models. Linköping University Electronic Press, 9 2022.

[8] H. Eldarzade, A. Quliyev, A. Sherifova, R. Shahmarova, and T. Xankishiyeva, "Research of the technology of opening holes in compact and powder materials," ETM - Equipment, Technologies, Materials, vol. 05, pp. 51–, 1 2021.

[9] M. Hossain and M. Alam, "Corporate social reporting (csr) and stakeholder accountability in bangladesh," International Journal of Accounting & Information Management, vol. 24, pp. 415–442, 10 2016.

[10] K. A. Rouf, "Concepts, theories, and approaches to management practices for sustainable businesses," International Journal of Research Studies in Management, vol. 7, 2 2018.

[11] E. L. Best and M. Binkley, "To be a b certified benefit corporation or not to be," The Contemporary Tax Journal, vol. 9, 7 2020.

[12] S. Taszharganov and T. Satkalieva, "Strategic priorities of the organization's reputation management development," SERIES OF SOCIAL AND HUMAN SCIENCES, vol. 1, pp. 161–168, 2 2019.

[13] Z. Cao, "Brand equity, warranty costs, and firm value," International Journal of Research in Marketing, vol. 39, no. 4, pp. 1166–1185, 2022.

[14] A. Mountfield, K. Hrajnoha, L. Koh, L. Lascenko, R. Puchala, and C. Schalch, Barriers to implementing sustainability experienced by middle managers in the fast-moving consumer goods and retail sector. Edward Elgar Publishing, 7 2021.

[15] M. Jell-Ojobor and A. Raha, "Being good at being good—the mediating role of an environmental management system in value-creating green supply chain management practices," Business Strategy and the Environment, vol. 31, pp. 1964–1984, 1 2022.

[16] B. Ackers, "Ethical considerations of corporate social responsibility - a south african perspective," South African Journal of Business Management, vol. 46, pp. 11–21, 3 2015.

[17] V. Kalra and K. S. Rana, Human African Trypanosomiasis, Sleeping Sickness, vol. 74, pp. 337–355. CRC Press, 1 2003.

[18] L. K. Stanislavská, L. Pilař, K. Margarisová, and R. Kvasnička, "Corporate social responsibility and social media: Comparison between developing and developed countries," Sustainability, vol. 12, pp. 5255–, 6 2020.

[19] R. C. Kim, "Rethinking corporate social responsibility under contemporary capitalism: Five ways to reinvent csr," Business Ethics, the Environment & Responsibility, vol. 31, pp. 346–362, 1 2022.

[20] E. A. da Rocha Garcia, I. A. Araujo, and A. R. A. Filho, "Disclosure social das empresas brasileiras referente aos objetivos de desenvolvimento sustentável da organização das nações unidas (onu): Um estudo da divulgação das metas relacionadas ao objetivo de igualdade de gênero," Contextus – Revista Contemporânea de Economia e Gestão, vol. 19, pp. 217–231, 7 2021.

[21] N. R. do Canto, A. C. D. Verschoore, P. Dias, and M. D. de Barcellos, "Corporate social responsibility and innovation: A multiple case study with brazilian companies," Revista de Gestão Social e Ambiental, vol. 14, pp. 93–113, 6 2020.

[22] R. Sharmelly and A. Klarin, "Customer value creation for the emerging market middle class: Perspectives from case studies in india," Journal of Risk and Financial Management, vol. 14, pp. 455–, 9 2021.

[23] M. Yang, J. Wang, and F. Yasmin, "Does higher business education championenvironmental sustainability for next generationof leaders? an assessment of in-school studentsand alumni's perspective," Polish Journal of Environmental Studies, vol. 30, pp. 5317–5332, 10 2021.

[24] E. Shekarian, B. Ijadi, A. Zare, and J. Majava, "Sustainable supply chain management: A comprehensive systematic review of industrial practices," Sustainability, vol. 14, pp. 7892–7892, 6 2022.

[25] J. Whitfield, L. A. Dioko, and D. E. Webber, "Scoring environmental credentials: a review of uk conference and meetings venues using the greener venue framework," Journal of Sustainable Tourism, vol. 22, pp. 299–318, 6 2013.

[26] A. John and V. P. J. Raj, "Employer brand and innovative work behaviour: Exploring the mediating role of employee engagement," Colombo Business Journal, vol. 11, pp. 93–113, 12 2020.

[27] Z. Cao and K. Xu, "Ceo narcissism, brand acquisition and disposal, and stock returns," Journal of the Academy of Marketing Science, vol. 50, no. 4, pp. 777–799, 2022.

[28] M. A. Camilleri, "A cost-benefit analysis of iso's standard on social responsibility: A review and appraisal," CSR, Sustainability, Ethics & Governance, pp. 159–175, 8 2018.

[29] M. Saayman and A. Giampiccoli, "Community-based and pro-poor tourism: Initial assessment of their relation to community development," European Journal of Tourism Research, vol. 12, pp. 145–190, 3 2016.

[30] D. Andulana, M. Calijan, and A. Albina, "Challenges and opportunities in philippine tourism amid the covid-19 pandemic," Recoletos Multidisciplinary Research Journal, vol. 9, pp. 67–87, 12 2021.

[31] J. Su, Y. Wei, S. Wang, and Q. Liu, "The impact of digital transformation on the total factor productivity of heavily polluting enterprises.," Scientific reports, vol. 13, pp. 6386–6386, 4 2023.

[32] M. Rossi, G. Festa, S. Chouaibi, M. Fait, and A. Papa, "The effects of business ethics and corporate social responsibility on intellectual capital voluntary disclosure," Journal of Intellectual Capital, vol. 22, pp. 1–23, 3 2021.

[33] S. Trevaskes and E. Nesossi, Control by Law, pp. 42–59. ANU Press, 6 2017.

[34] M. M. Ugolini, "Leveraging intersections in management theory and practice (extended abstracts)," Sinergie Italian Journal of Management, pp. 1–624, 10 2021.

[35] M. C. Oliveira, M. S. R. Júnior, S. H. D. O. Lima, and G. A. de Freitas, "The influence of the characteristics of the national business system in the disclosure of gender-related corporate social responsibility practices," Administrative Sciences, vol. 8, pp. 14–, 4 2018.

[36] L. Chen, Sustainability and company performance : Evidence from the manufacturing industry - Sustainability and company performance : Evidence from the manufacturing industry. Linköping University Electronic Press, 9 2015.

[37] P. M. S. Qureshi, New Advances in the Dairy Industry. 11 2021.

[38] J. D. Upstill-Goddard, J. Glass, A. R. Dainty, and I. Nicholson, "Implementing sustainability in small and medium-sized construction firms: The role of absorptive capacity," Engineering, Construction and Architectural Management, vol. 23, pp. 407–427, 7 2016.

[39] V. Bester and F. Cronjé, "The importance of a people–centred approach for corporate social responsibility : a case study of welverdiend and the surrounding community," The Journal for Transdisciplinary Research in Southern Africa, vol. 10, pp. 24–, 7 2014.

[40] S. Hamid, A. Saeed, U. Farooq, and F. Alnori, "A bibliometric retrospection of csr from the lens of finance and economics: Towards sustainable development," Sustainability, vol. 14, pp. 16852–16852, 12 2022.

[41] M. A. D. Mohammadi, A. Mardani, M. N. A. A. Khan, and D. Streimikiene, "Corporate sustainability disclosure and market valuation in a middle eastern nation: evidence from listed firms on the tehran stock exchange: sensitive industries versus non-sensitive industries," Economic Research-Ekonomska Istraživanja, vol. 31, pp. 1488–1511, 9 2018.

[42] P. S. Peek, Ancient Greek I - Module 30 - and , pp. 453–478. Open Book Publishers, 10 2021.

[43] W.-E. M. Ibrahim, "Societal responsibility of the foreign direct investment companies in the arabian," ,( ) vol. 39, pp. 227–251, 12 2019.

[44] M. G. E. Velter, "Journeying towards sustainable business models," 1 2022.

[45] W. Winit and S. Kantabutra, "Enhancing the prospect of corporate sustainability via brand equity: A stakeholder model," Sustainability, vol. 14, pp. 4998–4998, 4 2022.

[46] T. G. Hawkins, M. J. Gravier, and E. H. Powley, "Public versus private sector procurement ethics and strategy: what each sector can learn from the other," Strategic Direction, vol. 28, 3 2012.

[47] P. Gazzola, S. Amelio, and R. R. Pezzetti, "Csr as a driver of corporate reputation: Family firms in the italian luxury industry," International Journal of Business Administration, vol. 11, pp. 21–36, 10 2020.

[48] G. N. Prabhu and P. S. Aithal, "Quantitative abcd analysis of green banking practices and its impact on using green banking products," International Journal of Applied Engineering and Management Letters, pp. 28–66, 1 2023.

[49] S. Firk, A. Hanelt, J. Oehmichen, and M. Wolff, "Chief digital officers: An analysis of the presence of a centralized digital transformation role," Journal of Management Studies, vol. 58, pp. 1800–1831, 6 2021.