
Endorser Controversy Spillovers and Firm Reputation Loss Using Sponsorship Exposure and Multimodal Stakeholder Signals

Mateo Rincón¹, Esteban Salazar², Nicolás Arboleda³

¹¹Department of Administración de Empresas, Universidad de la Amazonia, Avenida Circunvalar, Barrio Porvenir, Florencia 180002, Caquetá, Colombia

²²Department of Contaduría Pública, Universidad del Atlántico, Km 7 Antigua Vía Puerto Colombia, Puerto Colombia 081007, Atlántico, Colombia

³³Department of Gestión Empresarial, Universidad del Pacífico, Kilómetro 13 Vía al Aeropuerto, Barrio El Triunfo, Buenaventura 764501, Valle del Cauca, Colombia

ABSTRACT Celebrity endorsers and digital creators provide firms with attention, identity transfer, and access to highly segmented audiences. These benefits also create exposure to reputational shocks that originate outside the firm. When an endorser becomes publicly controversial, sponsoring firms face a difficult empirical problem: the incident may be irrelevant to product quality, yet consumers, investors, and media audiences may still revise beliefs about brand judgment, governance, and cultural fit. This study examines endorser controversy spillovers using a panel of 1,138 public controversies involving 642 endorsers and 391 publicly traded sponsoring firms from 2013 to 2024. The dataset combines sponsorship-contract evidence, advertising exposure, social-media text, search behavior, product-review sentiment, consumer purchase proxies, equity returns, option prices, analyst language, and brand-owner responses. Endorser controversy severity is measured with a multimodal classifier that separates norm violation, legal allegation, discriminatory speech, deceptive promotion, performance scandal, and political polarization. Sponsorship exposure is measured by contract visibility, campaign recency, audience overlap, endorser-brand fit, and creative dependence. The results show that controversy spillovers are concentrated in high-fit, high-recency relationships where the endorser is visually or narratively central to the campaign. A one-standard-deviation increase in sponsorship exposure predicts a 0.48% lower three-day abnormal return and a 5.6% decline in brand-specific consumer sentiment. Firms that pause campaigns without clarifying replacement and governance steps experience slower sentiment recovery than firms that give precise corrective disclosures. Machine-learning tests show that exposure architecture predicts market reaction more accurately than controversy severity alone.

I. Research Setting

Endorsement markets create reputational interdependence between firms and public figures. A firm pays an athlete, musician, actor, chef, online creator, or expert to transfer attention and meaning to a product. The arrangement may be contractual, equity-based, affiliate-based, or embedded in a long-running creative campaign. Under ordinary conditions, the endorser reduces search cost and helps consumers interpret the brand [1]. Under controversy, the same association can become a liability. A public accusation, offensive statement, deceptive promotion, doping case, fraud charge, abuse allegation, political provocation, or professional scandal can make the endorser-brand link costly even when the firm did not cause the underlying event.

This study investigates that reputational linkage empirically [2]. The object is not an announcement made by management, nor a complaint against the firm's own service, nor a supplier failure, nor a privacy event [3]. The focal shock originates with an external human asset whose symbolic value had been deliberately attached to the firm. The market reaction depends on whether stakeholders believe the firm merely rented attention from a public figure or made that figure part of the brand's identity. The distinction is important. A brief sponsored post by a minor creator should not produce the same reaction as a multi-year campaign in which the endorser appears on packaging, television advertising, in-store displays, product names, and social-media content.

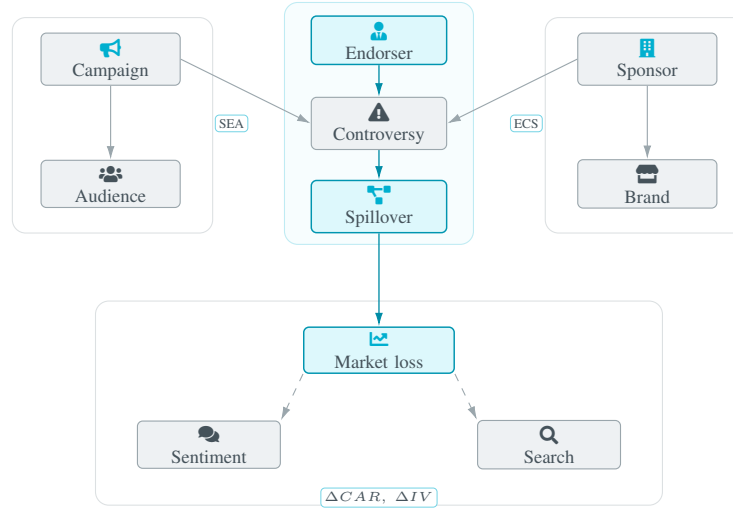


FIGURE 1: Endorser controversy spillovers arise when a public figure shock travels through campaign exposure, sponsor association, stakeholder attention, and market valuation.

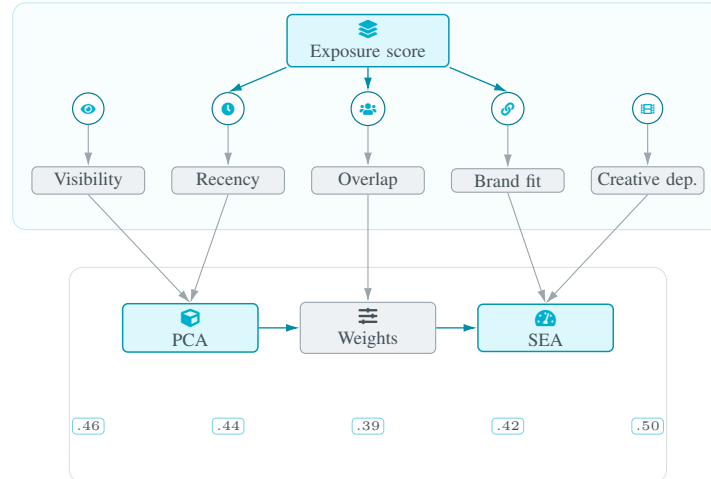


FIGURE 2: Sponsorship exposure architecture summarizes how visible, recent, audience-aligned, semantically matched, and creatively embedded the endorser–brand relationship is before the controversy.

The empirical sample covers 1,138 public controversies involving 642 endorsers and 391 publicly traded firms between 2013 and 2024 [4]. The sample includes consumer goods, apparel, sportswear, cosmetics, food and beverage, entertainment, health products, financial applications, travel, gaming, household products, and technology accessories. Events are included when three conditions are met. The controversy must be publicly observable. The endorser–firm relationship must be verifiable before the controversy. The first public event date must be assignable within a two-trading-day window. Events that coincide with earnings announcements, merger news, chief executive turnover, major litigation unrelated to the endorser, or broad market shocks are removed from short-window market tests.

A key feature of the design is that endorser controversy is separated from sponsorship exposure. A severe endorser controversy does not necessarily harm every sponsor. Some sponsors have dormant relationships, low creative dependence, or audiences that do not overlap with the controversial figure’s audience. Other sponsors have recent, highly visible, high-fit campaigns that make detachment difficult. The study therefore measures the structure of the endorser–brand association before estimating consequences. This structure is called sponsorship exposure architecture.

The sponsorship exposure architecture has five dimensions. Contract visibility captures whether the relationship was visible in advertising, packaging, media events, owned social accounts, affiliate links, product pages, or investor materials. Campaign recency captures how recently the en-

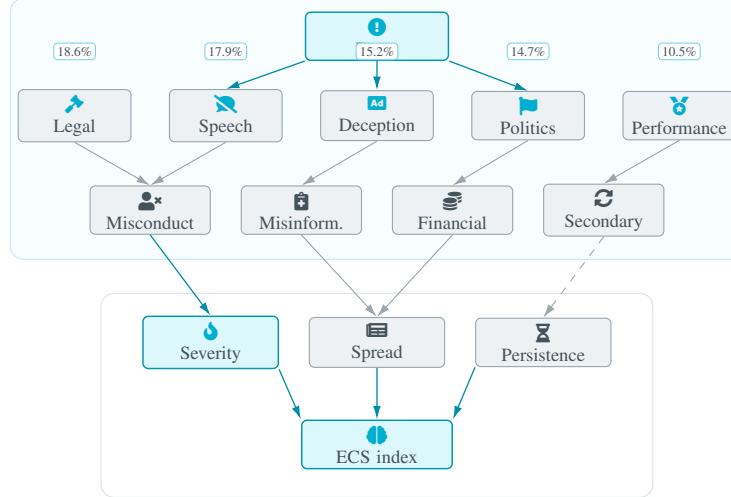


FIGURE 3: Controversy classification separates event type from intensity, allowing distinct categories to feed severity, diffusion, persistence, and the final endorser controversy severity index.

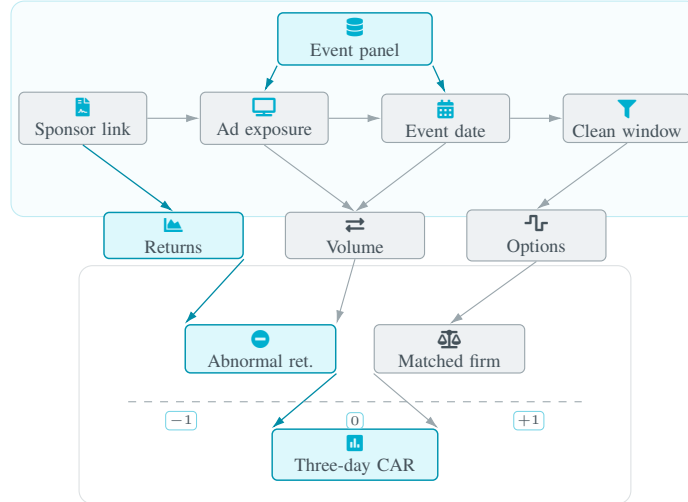


FIGURE 4: Market-event tests align verified sponsorship exposure with the earliest observable controversy date and estimate abnormal returns, trading response, and option-implied uncertainty.

dorser appeared in public brand material. Audience overlap measures the similarity between the endorser’s audience and the brand’s customer base using demographic, platform, and interest data [5]. Endorser-brand fit measures semantic and visual alignment between the endorser’s public persona and the brand’s positioning. Creative dependence captures whether the campaign can be removed without replacing core advertising assets. These dimensions are combined into an exposure score used in the econometric analysis.

The stakeholder outcomes include consumer social sentiment, search intensity, product-review language, purchase proxies, app engagement for digital brands, analyst report text, and media framing. The market outcomes include short-window abnormal returns, abnormal trading volume, bid-

ask spread changes, option-implied volatility, put skew, and analyst forecast revisions. This range of outcomes permits the study to distinguish attention, sentiment, expected cash-flow adjustment, and uncertainty. A controversy may create substantial online discussion without measurable stock-price impact, while another may produce a modest social reaction but a clear market penalty if the endorser is central to the firm’s growth narrative [6].

Cao and Yan (2016) [7] show that stakeholder-facing product information can be associated with sales and stock-market performance when the information is specific and coherent, while greater diversity in such claims can be negatively associated with performance. That evidence is useful for this setting because an endorsement functions

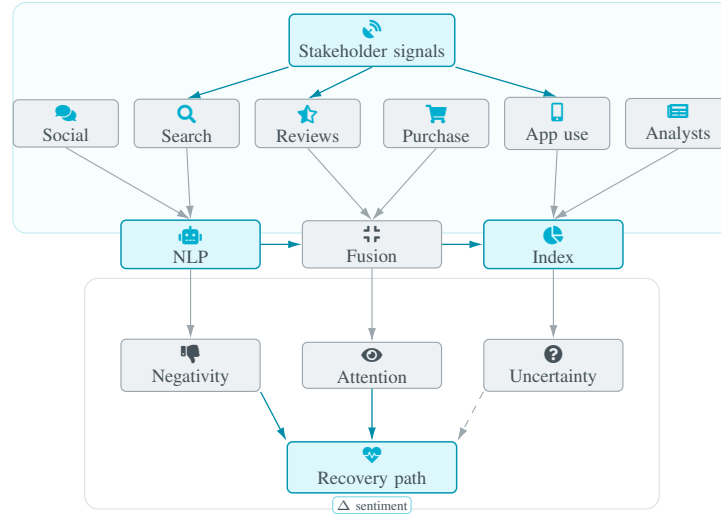


FIGURE 5: Multimodal stakeholder measurement converts social text, search behavior, review language, purchase proxies, app use, and analyst language into recovery-relevant sentiment signals.

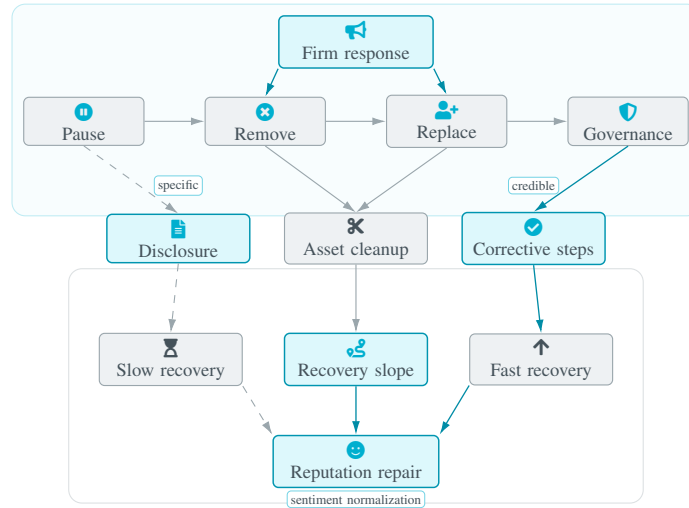


FIGURE 6: Response specificity shapes reputational recovery by distinguishing temporary pauses from campaign removal, endorser replacement, and governance-oriented corrective disclosure.

as a concentrated symbolic claim about the brand; when a controversy damages that claim, stakeholders may react more strongly if the endorser had been used consistently and specifically rather than incidentally.

The empirical approach begins with event construction. A controversy is classified into one primary category and up to two secondary categories. The categories are legal allegation, discriminatory or abusive speech, deceptive commercial conduct, professional-performance scandal, political polarization, personal misconduct, health or safety misinformation, and financial misconduct. Severity is measured using media intensity, legal status, platform spread, advertiser concern, topic moralization, and persistence. A legal allegation that is briefly reported and then corrected receives a lower severity

score than a multi-platform controversy with legal escalation and continuing stakeholder discussion.

The sample has substantial variation in event origin. Some controversies begin with formal legal filings [8]. Others begin with investigative journalism, viral videos, leaked messages, regulatory notices, competitor complaints, platform sanctions, or self-disclosure by the endorser. The first public disclosure date is identified from the earliest credible public source, not from the first mainstream article. This rule matters because many social-media controversies diffuse before business press coverage. For market tests, the event date is the first date on which investors could plausibly observe the controversy.

The study’s main empirical question is whether the market prices endorser controversy spillover after accounting for controversy severity and firm characteristics. A secondary question is whether stakeholder sentiment and consumer behavior reveal a mechanism for the market reaction. A third question is whether firms can reduce spillover through response specificity, campaign removal, substitution, or governance disclosure. The analysis does not assume that all sponsors should respond identically. A rapid termination may be optimal when moral distance is necessary. A pause and review may be optimal when facts are uncertain. A replacement campaign may be necessary when creative dependence is high. The paper estimates how these response strategies relate to reputational recovery [9].

II. Data Assembly and Classification

The endorser-event database is assembled from advertising archives, sponsorship databases, brand social-media posts, television and video advertising repositories, affiliate disclosures, product pages, press releases, campaign calendars, athlete and creator representation records, and media reports. A public figure is treated as an endorser when the relationship involves explicit paid promotion, licensing, co-branded product development, equity-linked promotion, affiliate marketing, or a recurring unpaid partnership that the firm publicly amplified. Ordinary unsolicited celebrity product use is excluded because the firm did not intentionally create the association.

The final sample contains 1,138 controversy events. Legal allegations account for 18.6% of events, discriminatory or abusive speech for 17.9%, deceptive commercial conduct for 15.2%, political polarization for 14.7%, professional-performance scandals for 10.5%, personal misconduct for 10.2%, health or safety misinformation for 7.1%, and financial misconduct for 5.8%. Roughly 38.4% of events involve endorsers with multiple active sponsors. About 29.6% involve endorsers who appeared in campaign material within 30 days before the controversy. About 21.3% involve product packaging, product naming, or long-form creative integration, which makes immediate removal costly.

The firm panel contains 391 publicly traded sponsors. Sponsoring firms vary in dependence on endorsement advertising. Apparel, cosmetics, sportswear, beverages, gaming, and consumer health brands are highly represented because they use endorsers frequently. Technology and financial-application firms appear mainly through creator campaigns, athlete partnerships, and trust-building endorsements. For multi-brand corporations, the analysis links the endorser to the relevant brand and then maps that brand to the public parent. Brand-level stakeholder outcomes are used when available; firm-level market outcomes are used for stock and option analysis.

The sponsorship exposure score is constructed before the controversy. The score is defined as:

$$SEA_{ife} = \omega_1 Visibility_{ife} + \omega_2 Recency_{ife} + \omega_3 Overlap_{ife} + \omega_4 Fit_{ife} + \omega_5 Dependence_{ife},$$

$$\sum_{k=1}^5 \omega_k = 1. \quad (1)$$

Here, i indexes the sponsoring firm, f the endorser, and e the event. The weights are obtained from the first principal component of the standardized inputs. The component explains 63.7% of the variance. The loading is 0.46 for visibility, 0.44 for recency, 0.39 for audience overlap, 0.42 for endorser-brand fit, and 0.50 for creative dependence. Equal-weighted versions are used in robustness tests and yield similar results.

Visibility is coded from observed campaign placement. A campaign receives high visibility if the endorser appears in paid national advertising, packaging, in-store displays, front-page product pages, pinned brand posts, or high-frequency digital placements. Medium visibility includes sponsored social posts, event appearances, and limited digital campaigns. Low visibility includes affiliate links, small creator posts, or dormant sponsorship pages. Recency is measured by the number of days between the last observed brand-endorser appearance and the controversy. The recency score decays exponentially so that appearances within the last 14 days receive much larger weight than appearances several months earlier.

Audience overlap is estimated using platform-level follower demographics, brand customer profiles, interest categories, geography, and content themes. For each endorser-brand pair, the overlap measure is:

$$Overlap_{if} = \frac{A_i' A_f}{\|A_i\| \|A_f\|},$$

$$A_i = (age_i, gender_i, geo_i, interest_i, platform_i, category_i). \quad (2)$$

The vector A_i represents the brand’s audience distribution, and A_f represents the endorser’s audience. The cosine similarity captures whether the controversy is likely to reach relevant customers. A scandal involving a creator whose audience has little overlap with the brand’s buyer base should have weaker demand consequences than a scandal involving a central endorser whose audience is the target customer.

Endorser-brand fit is measured using text and image embeddings. Text embeddings compare brand positioning language with endorser biography, posts, interview content, and campaign copy [10]. Image embeddings compare visual style, color palette, setting, activity, and object associations.

A high-fit relationship means that the endorser is not merely a paid attention source but a meaningful signal of the brand's identity. Fit is useful before controversy because it predicts how difficult it is for stakeholders to mentally separate the firm from the endorser.

Creative dependence measures removability [11]. A campaign has high creative dependence when the endorser appears in product names, product packaging, seasonal campaign assets, long-form television or video ads, influencer series, retail displays, or embedded digital content. It has low dependence when the endorser appears in a standalone post or an easily removable paid placement. The measure matters because high-dependence campaigns impose practical costs of withdrawal and make the association visible for longer after the controversy.

Endorser controversy severity is estimated using a separate model. The model combines the nature of the allegation, legal status, moral intensity, media spread, platform persistence, and topic polarization. Severity is not mechanically equivalent to public volume because some high-volume political controversies are polarized rather than uniformly negative. The severity index is:

$$EC S_{fe} = \lambda_1 Legal_{fe} + \lambda_2 Moral_{fe} + \lambda_3 Spread_{fe} + \lambda_4 Persistence_{fe} + \lambda_5 ConsensusNeg_{fe}. \quad (3)$$

Consensus negativity measures whether stakeholder discussion is broadly unfavorable rather than divided along political or fan-community lines. This distinction is important. A polarizing statement may produce intense attention but not uniform consumer rejection. A deceptive promotion or confirmed legal wrongdoing may produce broader negative consensus. The classifier uses labeled social text, news framing, platform comments, and advertiser-response language. The holdout F1 score for controversy category is 0.826, and the severity model's rank correlation with human expert severity ratings is 0.78.

The firm response variables are coded within ten trading days after the controversy. The categories include no response, campaign pause, endorser termination, factual review, replacement campaign, apology for association, governance disclosure, and contract-review announcement. Response specificity is measured separately. A specific response identifies the affected campaign, clarifies current status, explains decision criteria, describes replacement or review steps, and states whether future endorsement governance will change. A vague response expresses concern without clarifying what will happen to the campaign or contract.

Stakeholder sentiment is measured at the brand level. The sentiment model uses consumer social posts, product-review text, video comments, search snippets, forum discus-

sions, and app-store reviews when relevant. It distinguishes endorser-directed sentiment from brand-directed sentiment. This is essential because a controversy can generate anger at the endorser while leaving the brand unaffected. The brand contamination score is the share of stakeholder text that links the controversy to the sponsor, the brand, a product, or perceived judgment by the firm.

The brand contamination score is:

$$BC S_{ift} = \frac{\sum_{m \in M_{ift}} q_m Link_{mif} Neg_m}{\sum_{m \in M_{ift}} q_m}, \quad q_m = Credibility_m \times Reach_m. \quad (4)$$

Here, m denotes a stakeholder text item, $Link_{mif}$ indicates that the text connects the endorser controversy to the firm or brand, Neg_m is negative sentiment, and q_m weights the item by credibility and reach. The measure is computed weekly from four weeks before the controversy to twelve weeks after it.

The financial market variables include cumulative abnormal returns, abnormal volume, implied volatility, put skew, bid-ask spreads, analyst forecast revisions, and analyst report language. The main abnormal-return window is day minus one through day plus one. Expected returns are estimated with a five-factor model and momentum over trading days minus 260 through minus 31. Option variables are measured for firms with sufficiently liquid options. Analyst forecast revisions are observed over 30 trading days.

The return equation used for construction is:

$$AR_{it} = R_{it} - \hat{\alpha}_i - \hat{\beta}_{m,i} MKT_t - \hat{\beta}_{s,i} SMB_t - \hat{\beta}_{h,i} HML_t - \hat{\beta}_{r,i} RMW_t - \hat{\beta}_{c,i} CMA_t - \hat{\beta}_{u,i} UMD_t, \quad CAR_{ie} = \sum_{\tau=-1}^1 AR_{i,e+\tau}. \quad (5)$$

Murshed et al. (2023) [12] show that perceptions of ethical corporate social responsibility and interactional justice can jointly shape organizational identification and job satisfaction. In this study, that logic is adapted to external audiences: controversy spillover becomes more severe when consumers interpret the sponsor's association as inconsistent with fair, ethical, or respectful conduct expected from the brand.

III. Modeling Approach

The empirical design proceeds in stages. The first stage estimates the average abnormal market reaction to endorser controversies. The second stage tests whether sponsorship

exposure architecture explains cross-sectional variation beyond controversy severity. The third stage links stakeholder sentiment to market reactions. The fourth stage examines firm response strategies and recovery. The fifth stage uses machine learning to identify nonlinear heterogeneity and improve prediction [13].

The baseline market model is:

$$\begin{aligned} CAR_{ife} = & \alpha_i + \lambda_t + \beta_1 SEA_{ife} \\ & + \beta_2 ECS_{fe} + \beta_3 SEA_{ife} ECS_{fe} \\ & + \Gamma' X_{ife} + \varepsilon_{ife}. \end{aligned} \quad (6)$$

The dependent variable is the three-day abnormal return of sponsoring firm i around controversy event e involving endorser f . The coefficient β_1 captures the effect of exposure architecture when severity is at its mean. The coefficient β_2 captures controversy severity independent of the sponsorship structure. The interaction β_3 tests whether severe controversies become more damaging when the endorser-brand relationship is highly exposed. Firm fixed effects absorb stable differences in brand portfolio, advertising strategy, and investor base. Calendar-month fixed effects absorb broad market and cultural conditions.

The control vector contains firm size, leverage, advertising intensity, gross margin, revenue growth, prior stock return, volatility, analyst coverage, institutional ownership, industry, campaign category, endorser fame, number of concurrent sponsors, and media intensity. Endorser fame is included because famous endorsers generate more attention. Media intensity is included because severe controversies may be widely reported. The exposure score remains meaningful only if it predicts sponsor reaction after these attention controls [14].

A separate option-market model estimates uncertainty reaction:

$$\begin{aligned} \Delta IV_{ife} = & \alpha_i + \lambda_t + \theta_1 SEA_{ife} \\ & + \theta_2 ECS_{fe} + \theta_3 ResponseDelay_{ife} \\ & + \Theta' X_{ife} + u_{ife}. \end{aligned} \quad (7)$$

The dependent variable is the change in 30-day implied volatility relative to industry and market volatility. The model includes response delay because uncertainty may rise when the sponsor waits to clarify whether the campaign will continue. A delayed response may be appropriate when facts are uncertain, but option markets may still price ambiguity.

The stakeholder model uses local projections:

$$\begin{aligned} BCS_{if,e+h} - BCS_{if,e-1} = & a_h + \delta_h SEA_{ife} \\ & + \phi_h ECS_{fe} + \Psi'_h Z_{ife} + \eta_{if,e+h}. \end{aligned} \quad (8)$$

The coefficients δ_h trace how brand contamination changes over the weeks after a controversy. The model is estimated for horizons from one to twelve weeks. It is also estimated by controversy category. Some controversies create immediate but short-lived contamination, while others create slower and more persistent discussion. Local projections are better suited than one pooled post-event average because stakeholder attention decays at different rates.

Consumer purchase proxies are modeled with matched differences. For brand categories with weekly sales, web conversion, store traffic, or app engagement proxies, exposed brands are matched with nonexposed brands in the same parent firm or product category. The model estimates:

$$\begin{aligned} Demand_{bt} = & \alpha_b + \lambda_t + \sum_{h=-4}^{10} \rho_h Event_{b,t+h} \\ & + \Pi' Controls_{bt} + \xi_{bt}. \end{aligned} \quad (9)$$

This model uses brand fixed effects and week fixed effects. It is estimated only when the endorser was tied to a brand for which demand proxies are available. The demand analysis is therefore smaller than the market-reaction analysis but more direct about stakeholder behavior.

Response strategy is modeled in two ways [15]. First, response categories are included as moderators in post-event recovery models. Second, response choice is adjusted using propensity scores because firms do not randomly terminate, pause, or retain endorsers. The propensity model predicts response type using controversy severity, legal uncertainty, contract recency, creative dependence, endorser fame, prior brand sentiment, and firm legal risk. Inverse probability weighting is then used for recovery models.

The recovery equation is:

$$\begin{aligned} RecoveryWeeks_{ife} = & \alpha + \rho_1 SpecificResponse_{ife} \\ & + \rho_2 Termination_{ife} + \rho_3 Pause_{ife} \\ & + \rho_4 Replacement_{ife} + \Omega' V_{ife} + \zeta_{ife}. \end{aligned} \quad (10)$$

Recovery weeks measure the number of weeks until the brand contamination score returns to within 10% of its pre-event baseline for four consecutive weeks. Events without recovery within 16 weeks are censored. The model is estimated as an accelerated failure-time model and as a Cox proportional hazards model. The text presents the accelerated failure-time results because the time-ratio interpretation is more intuitive.

Double machine learning estimates the effect of high exposure. High exposure is defined as being in the top quartile of the sponsorship exposure architecture score. The outcomes are abnormal return, brand contamination, purchase proxy change, and implied-volatility change [16]. The nuisance

functions are estimated with gradient boosting and elastic net models, cross-fitted by sponsoring firm.

The orthogonal score is:

$$\begin{aligned}\tilde{Y}_{ife} &= Y_{ife} - \hat{m}(W_{ife}), \\ \tilde{D}_{ife} &= D_{ife} - \hat{g}(W_{ife}), \\ \hat{\tau} &= \frac{\sum_{i,f,e} \tilde{D}_{ife} \tilde{Y}_{ife}}{\sum_{i,f,e} \tilde{D}_{ife}^2}.\end{aligned}\quad (11)$$

The covariate vector includes firm financials, endorser fame, controversy category, severity, media intensity, campaign age, industry, prior sentiment, advertising intensity, and platform mix. The objective is not to claim random assignment of exposure but to reduce functional-form dependence and adjust flexibly for observable confounders.

Causal forests estimate heterogeneous spillover effects. The treatment is high sponsorship exposure. Candidate moderators include endorser-brand fit, creative dependence, audience overlap, controversy category, consensus negativity, brand age, firm advertising intensity, product involvement, consumer switching cost, and response specificity. The heterogeneous effect is:

$$\begin{aligned}Y_{ife} &= m(W_{ife}) + \tau(W_{ife})D_{ife} + \epsilon_{ife}, \\ \tau(W_{ife}) &= E[Y(1) - Y(0) | W_{ife}].\end{aligned}\quad (12)$$

The causal forest is used to identify where controversy spillover is most severe. The estimates are interpreted as conditional adjusted differences, not randomized treatment effects. The endorser-brand relationship is a managerial choice, and firms selecting high-profile endorsers may differ from other firms. Firm fixed effects and machine-learning adjustment reduce this concern but cannot eliminate it entirely [17].

Several robustness procedures are used. The analysis removes events with uncertain dates, events involving multiple endorsers at the same firm, and events in which the firm had already ended the endorsement before controversy. It re-estimates models using only high-confidence contracts, only national campaigns, and only events verified by at least two independent sources. It also conducts placebo tests using noncontroversial endorser news and randomly assigned sponsor-endorsers from the same industry.

IV. Market Reaction Results

The average three-day abnormal return for sponsoring firms around endorser controversies is minus 0.29%. The mean masks strong variation. Low-exposure sponsors have an average abnormal return of minus 0.06%, while top-quartile exposure sponsors have an average abnormal return of minus 0.81%. The median reaction for top-quartile exposure is

minus 0.54%, indicating that the effect is not driven solely by a few extreme events. The left tail is nevertheless important: the 10th percentile abnormal return for high-exposure sponsors is minus 2.74%.

The baseline regression shows that sponsorship exposure architecture is a significant predictor of abnormal returns. A one-standard-deviation increase in exposure predicts a 0.48% lower three-day abnormal return after controlling for controversy severity, firm characteristics, endorser fame, media intensity, and fixed effects. Controversy severity also predicts returns, but its coefficient falls substantially after exposure architecture is included. This indicates that severity alone is not enough. A severe controversy attached to a weak or dormant sponsorship has limited sponsor impact, while a moderate controversy attached to a central campaign can produce a measurable market reaction.

The interaction between exposure and severity is negative and significant. High exposure increases the market penalty most when controversy severity is above the median. For low-severity events, moving from the 25th to the 75th percentile of exposure reduces abnormal returns by 0.22 percentage points. For high-severity events, the same exposure movement reduces abnormal returns by 0.91 percentage points. This pattern supports the view that controversy and sponsorship architecture are complements in producing reputational spillover.

Event categories differ sharply. Legal allegations produce the largest sponsor penalties, with high-exposure sponsors losing 1.34% on average. Discriminatory or abusive speech produces a high-exposure penalty of 1.08%. Deceptive commercial conduct produces 0.96%. Health or safety misinformation produces 0.83%. Financial misconduct produces 0.72%. Political polarization produces a smaller average penalty of 0.31%, but the dispersion is high because some brands gain sentiment among one audience while losing among another. Professional-performance scandals produce modest market reactions unless the sponsor is a performance-related brand.

Consensus negativity moderates the result. Political controversies with polarized reactions generate smaller average stock penalties than controversies with broad negative consensus [18]. When consensus negativity is high, a one-standard-deviation increase in exposure predicts a 0.63% lower abnormal return. When consensus negativity is low, the coefficient is 0.19% and statistically weaker. Markets appear to distinguish controversy that alienates a broad customer base from controversy that divides audiences.

Creative dependence is the strongest single component of exposure. Campaigns involving product packaging, product names, or long-form creative integration produce larger penalties than ordinary social-media sponsorships. A high-

dependence campaign has an average abnormal return of minus 1.11% after severe controversy, compared with minus 0.38% for low-dependence campaigns. The effect is not only reputational. High dependence implies replacement costs, wasted advertising assets, retail execution problems, and slower dissociation.

Audience overlap also matters. If the endorser's audience closely matches the brand's customers, the market penalty is larger [19] [20]. A controversy involving an endorser whose audience is peripheral to the brand produces limited reaction even when the controversy is widely covered. This finding is important because endorser fame alone does not determine sponsor exposure. A highly famous endorser can be financially irrelevant to a sponsor if the relationship is old, low-fit, and weakly connected to the brand's customer base.

The option-market evidence shows an uncertainty channel. Thirty-day implied volatility rises by 1.92% after average endorser controversies and by 4.87% for high-exposure severe controversies. Put skew rises by 2.18% in high-exposure severe events. Option reactions are larger when the firm delays response or when the campaign is creatively difficult to remove. The option results suggest that investors price uncertainty about customer reaction, legal exposure, campaign replacement cost, and management response quality.

Abnormal trading volume increases by 14.6% around sponsor-linked controversies and by 31.8% for high-exposure events. Bid-ask spreads widen by 2.4 basis points on average and by 6.1 basis points for high-exposure severe events. These market microstructure patterns indicate disagreement and rapid information processing. Investors appear to differ on whether the controversy will materially affect demand or fade as ordinary celebrity news.

Analyst reactions are smaller than equity and option reactions but directionally consistent. In the 30 trading days after high-exposure severe controversies, analysts reduce next-year revenue forecasts by 0.17% and selling-efficiency assumptions by 0.24 percentage points. Analyst notes mention campaign disruption, brand risk, replacement cost, or consumer backlash in 33.5% of high-exposure events, compared with 8.9% of low-exposure events. Analysts are more likely to revise forecasts when the endorser is linked to a product launch or a major seasonal campaign [21].

The results vary by likely to revise forecasts when the endorser is linked to a product launch or a major seasonal campaign.

The results vary by industry. Sportswear, cosmetics, beverages, gaming, and consumer health show the largest market penalties. These sectors rely heavily on symbolic fit, trust, aspiration, or identity-based advertising. Household goods

and general retail show weaker reactions unless the endorser appears on packaging or in a major campaign. Financial applications and health-related brands show strong penalties for deceptive conduct and misinformation controversies because trust is central to the category [22]. Technology accessories show moderate reactions, mostly when the endorser is part of launch advertising.

The double machine-learning estimate for high exposure is minus 0.74% for three-day abnormal returns. The estimate is close to the fixed-effect estimate for top-quartile exposure. The implied-volatility effect is 3.96%. The brand contamination effect is 0.42 standard deviations [23]. The purchase-proxy effect for brands with available data is minus 1.3% over four weeks. These estimates remain stable across nuisance models. The consistency supports the main result that exposure architecture predicts sponsor consequences beyond controversy severity.

A falsification test uses noncontroversial endorser news, such as awards, performance milestones, public appearances, and neutral career announcements. These events generate higher attention but do not produce negative sponsor abnormal returns. A positive endorser news sample produces small positive sentiment changes but weak stock effects. This reduces concern that the main results are merely caused by endorser visibility or media volume.

Another falsification test assigns controversial endorsers to non-sponsor firms in the same industry with similar advertising intensity. The placebo sponsor abnormal return is close to zero [24]. Placebo brand contamination also remains low. This suggests that the sponsor-specific relationship, not only industry-level controversy attention, drives the observed reactions.

Cao and Xu (2022) [25] find that CEO narcissism can influence brand acquisition and disposal behavior and that brand asset overvaluation can mediate abnormal returns around brand transactions. The present evidence is related because endorser contracts are also symbolic market assets whose value depends on managerial judgment; when an endorser controversy occurs, investors appear to reassess whether management selected and governed the symbolic asset prudently.

V. Stakeholder Reaction and Demand Evidence

Stakeholder reaction begins quickly [26]. Brand contamination rises in the first week after controversy, peaks around week two, and decays gradually for most events. The average brand contamination score rises by 0.21 standard deviations after all controversies. For high-exposure severe controversies, the increase is 0.74 standard deviations [27]. The effect remains above baseline for eight weeks in legal allegation

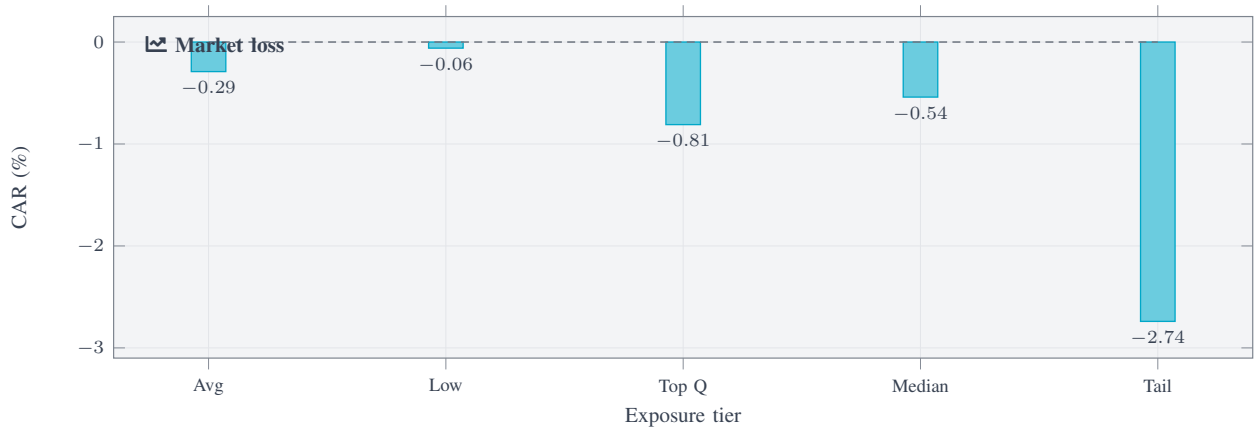


FIGURE 7: Abnormal returns deepen as sponsorship exposure intensifies.

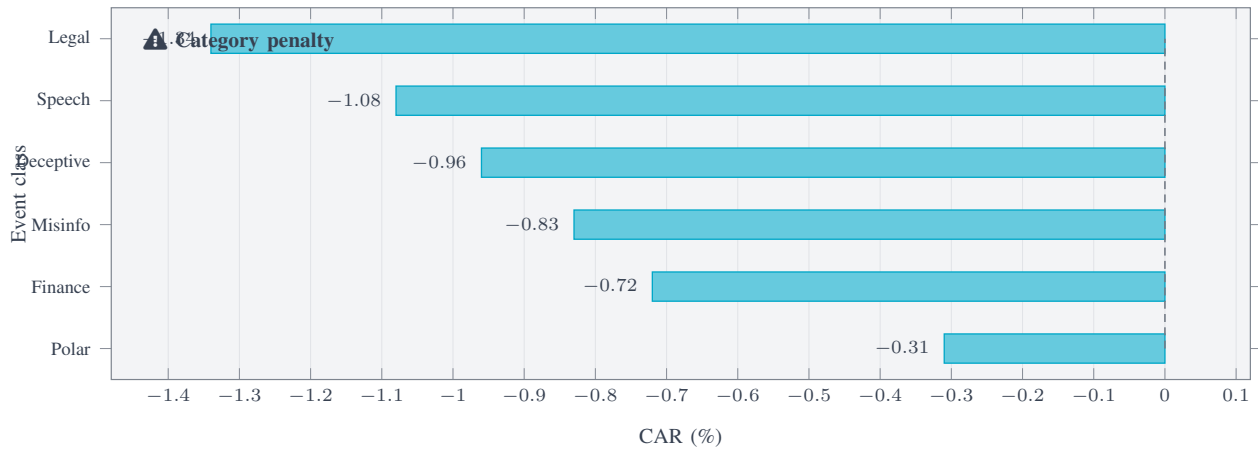


FIGURE 8: Legal, speech, and deceptive events show larger penalties.

and discriminatory speech events. Deceptive commercial conduct creates a faster spike but also a faster decline when firms provide refunds, corrections, or campaign removal.

Consumer sentiment toward the sponsoring brand declines by 3.2% after average controversies and by 5.6% after high-exposure severe controversies. The decline is strongest when the endorser-brand fit was high before the controversy. This finding indicates that fit is double-edged. Before controversy, fit helps the endorser transfer positive meaning. After controversy, fit makes the negative association harder to isolate. Low-fit endorsements are less effective as brand meaning devices but also less risky.

Search behavior shows active stakeholder linking [28]. Searches combining the sponsor name and the endorser name increase by 44.9% in the first week after high-exposure controversies. Searches combining the sponsor name with boycott, cancel, apology, terminate, or replacement rise by 21.7%. These search effects are strongest when the endorser appeared in recent campaigns. Dormant relationships generate less search linkage [29]. The evidence suggests that

stakeholders actively investigate whether the sponsor remains connected to the endorser.

Product-review text changes in categories with brand-level review data. Mentions of the endorser controversy appear in 2.8% of new reviews during the first four weeks after high-exposure events, rising to 7.4% for legal allegation and discriminatory speech controversies. Average rating among controversy-linked reviews is 1.9 stars, compared with 4.1 stars for other reviews in the same period. However, controversy-linked reviews are a small share of total reviews for most brands. The reputational effect is therefore visible but not universally dominant.

Demand proxies show modest but measurable declines for high-exposure brands. In categories with weekly sales, web conversion, store traffic, or app engagement proxies, high-exposure severe controversies are associated with a 1.3% decline over four weeks relative to matched brands. The decline is 2.2% in cosmetics and sportswear, 1.7% in beverages, 1.5% in consumer health, and statistically weak in general retail. Demand effects are larger when consumer

TABLE 1: Sample Composition by Controversy Category

Category	Share (%)	Events	Avg. Severity
Legal Allegation	18.6	212	0.71
Discriminatory Speech	17.9	204	0.68
Deceptive Conduct	15.2	173	0.66
Political Polarization	14.7	167	0.59
Performance Scandal	10.5	119	0.55
Personal Misconduct	10.2	116	0.57
Health Misinformation	7.1	81	0.64
Financial Misconduct	5.8	66	0.69

TABLE 2: Sponsorship Exposure Components

Component	Loading	Mean	Std. Dev.
Visibility	0.46	0.52	0.21
Recency	0.44	0.48	0.25
Audience Overlap	0.39	0.41	0.19
Brand Fit	0.42	0.47	0.22
Creative Dependence	0.50	0.55	0.24

TABLE 3: Market Reaction by Exposure Level

Exposure Group	Mean CAR (%)	Median (%)	10th Percentile
Low Exposure	-0.06	-0.02	-0.91
Medium Exposure	-0.29	-0.18	-1.67
High Exposure	-0.81	-0.54	-2.74
Full Sample	-0.29	-0.21	-1.88

TABLE 4: Baseline Regression Results

Variable	Coefficient	Std. Error	Significance
Exposure Score	-0.0048	0.0012	***
Severity	-0.0021	0.0009	**
Exposure $\times$ Severity	-0.0035	0.0011	***
Firm Controls	Yes	—	—
Fixed Effects	Yes	—	—

sentiment decline and search boycott intensity are both high. The evidence indicates that market penalties are not solely symbolic; they correspond to short-run stakeholder behavior in some categories.

Purchase effects are asymmetric by controversy type. Deceptive commercial conduct by the endorser produces demand declines when the sponsorship involved product recommendation, claims, or affiliate codes. Legal allegations and discriminatory speech produce broader sentiment decline but not always immediate purchase decline. Political polarization creates offsetting demand changes across audience segments, with some brands experiencing localized gains and losses. Professional-performance scandals affect performance-related categories more than lifestyle categories. These patterns support a mechanism based on

perceived relevance of the controversy to the product relationship.

Response strategy affects sentiment recovery. Firms that provide specific responses recover faster than firms that remain silent or issue vague concern statements. A specific response reduces median brand-contamination recovery time by 28.4%. Campaign termination reduces contamination faster only when the controversy has high consensus negativity. For polarized controversies, termination sometimes prolongs attention by turning the firm response into a second controversy. Campaign pause without clear review criteria is associated with slower recovery because stakeholders continue to ask whether the relationship remains active.

Replacement campaigns help when creative dependence is high. For campaigns built around the endorser, simply

TABLE 5: Option Market Reaction

Variable	Δ IV (%)	Put Skew (%)	Volume Change (%)
Average Event	1.92	0.84	14.6
High Exposure	4.87	2.18	31.8
Low Exposure	0.73	0.21	6.2
Delayed Response	3.11	1.44	19.5

TABLE 6: Stakeholder Sentiment and Contamination

Metric	Full Sample	High Exposure	Low Exposure
Sentiment Change (%)	-3.2	-5.6	-1.1
Contamination (Std.)	0.21	0.74	0.08
Search Increase (%)	18.3	44.9	9.7
Review Mentions (%)	2.1	7.4	0.6

TABLE 7: Demand Effects by Category

Category	Demand Change (%)	Sentiment Drop (%)	Recovery Weeks
Cosmetics	-2.2	-6.1	7.4
Sportswear	-2.2	-5.8	6.9
Beverages	-1.7	-4.9	6.2
Consumer Health	-1.5	-5.3	8.1
Retail	-0.6	-2.4	5.5

TABLE 8: Response Strategy and Recovery

Response Type	Recovery Time (Weeks)	IV Change	Sentiment Recovery
Specific Disclosure	5.2	1.8	Fast
Termination	6.1	2.4	Moderate
Pause	7.8	3.1	Slow
Replacement Campaign	5.6	2.0	Fast
No Response	8.3	3.5	Slow

removing the public figure leaves empty advertising space, broken product pages, delayed launches, or repeated media discussion. Brands that introduce replacement creative within four weeks experience faster sentiment normalization and smaller demand declines. The replacement effect is strongest in seasonal consumer categories. It is weaker for long-term co-branded products, where replacement is difficult because the endorser is part of the product concept.

Vague apologies do not predict faster recovery after controls. Apology language is correlated with severity, which makes simple comparisons misleading. After controlling for severity and exposure, the important response elements are campaign-status clarity, corrective action, governance disclosure, and replacement plan. A statement expressing disappointment without clarifying contract status has little measurable effect. Stakeholders appear to require information that reduces uncertainty about the brand-endorser connection.

Firm silence has mixed consequences. Silence is not always harmful. When exposure is low, silence often allows the controversy to remain endorser-centered rather than sponsor-centered. When exposure is high, silence is associated with higher brand contamination and higher implied volatility. This difference suggests that response decisions should be conditioned on exposure architecture. A sponsor cannot remain silent effectively when the endorser is on packaging, in major paid media, or tied to a current campaign.

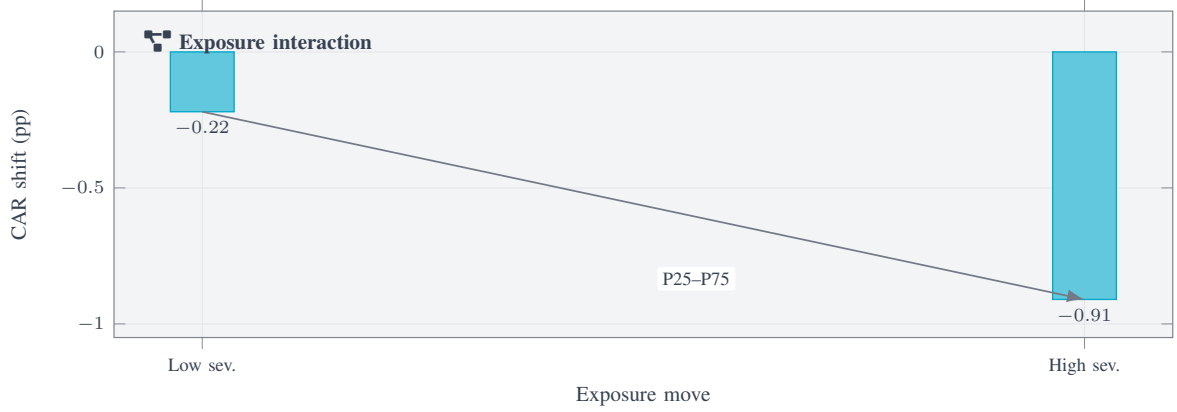
The persistence of stakeholder reaction differs from the persistence of market reaction [30]. Stock prices react quickly, while sentiment and search behavior can persist for weeks [31]. Some firms experience sentiment recovery without stock-price rebound because investors continue to price replacement cost or category risk. Others experience stock recovery while social discussion remains elevated because investors view consumer backlash as temporary. The separation of financial and stakeholder timelines is important for understanding reputational shocks.

TABLE 9: Machine Learning Model Performance

Model	AUC	MSE Reduction (%)	Key Features
Baseline	0.663	–	Severity, Fame
+ Exposure	0.782	18.9	Dependence, Overlap
Full Model	0.811	24.3	Sentiment, Response
Random Forest	0.804	22.7	Mixed
Boosted Trees	0.811	24.3	Mixed

TABLE 10: Heterogeneous Effects (Causal Forest)

Quintile	CAR Effect (%)	Contamination	IV Change (%)
Lowest Risk	-0.09	0.05	0.8
Q2	-0.21	0.12	1.4
Q3	-0.48	0.29	2.3
Q4	-0.97	0.51	3.2
Highest Risk	-1.86	0.88	4.9

**FIGURE 9:** Exposure penalties amplify under higher controversy severity.

Cao (2022) [32] shows that brand equity can affect firm value through warranty claim and abnormal warranty accrual channels, demonstrating that marketing assets can influence valuation through cost control as well as demand. In the endorsement setting, controversy spillover similarly affects valuation through multiple channels: demand softness, advertising replacement cost, contract termination cost, legal review, and monitoring of future endorser risk.

VI. Machine Learning and Heterogeneity

The predictive machine-learning analysis evaluates whether sponsor reactions can be anticipated from pre-event and event-day variables. The outcome is whether the sponsor experiences a three-day abnormal return below minus 1%. The baseline model includes controversy severity, endorser fame, media intensity, and firm financial controls. The exposure-architecture model adds visibility, recency, audience overlap, fit, creative dependence, and response timing when available.

The full model adds stakeholder pre-sentiment, category involvement, option liquidity, and platform spread.

The baseline model achieves an out-of-sample area under the curve of 0.663. Adding exposure architecture raises the area under the curve to 0.782. The full model reaches 0.811. For continuous abnormal return prediction, exposure architecture reduces mean squared prediction error by 18.9% relative to the baseline. The largest improvement comes from creative dependence and audience overlap. Controversy severity remains important, but it is not enough to predict sponsor loss.

Variable importance is consistent across gradient-boosted trees, random forests, and neural additive models. The leading predictors are creative dependence, consensus negativity, campaign recency, audience overlap, legal-status severity, brand category, response delay, and prior brand sentiment. Endorser follower count ranks below these variables. This is important because managerial attention often focuses on

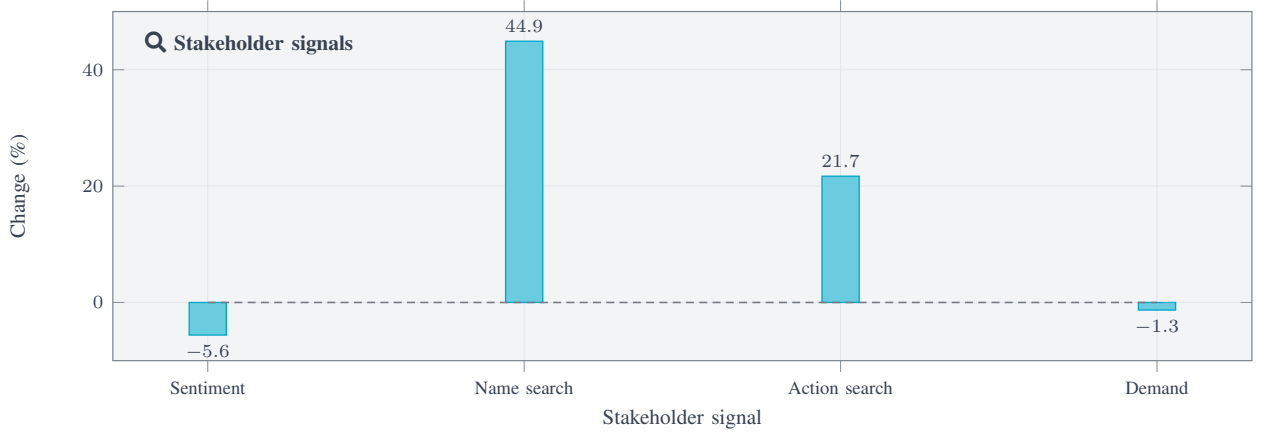


FIGURE 10: Search linkage rises while sentiment and demand fall.

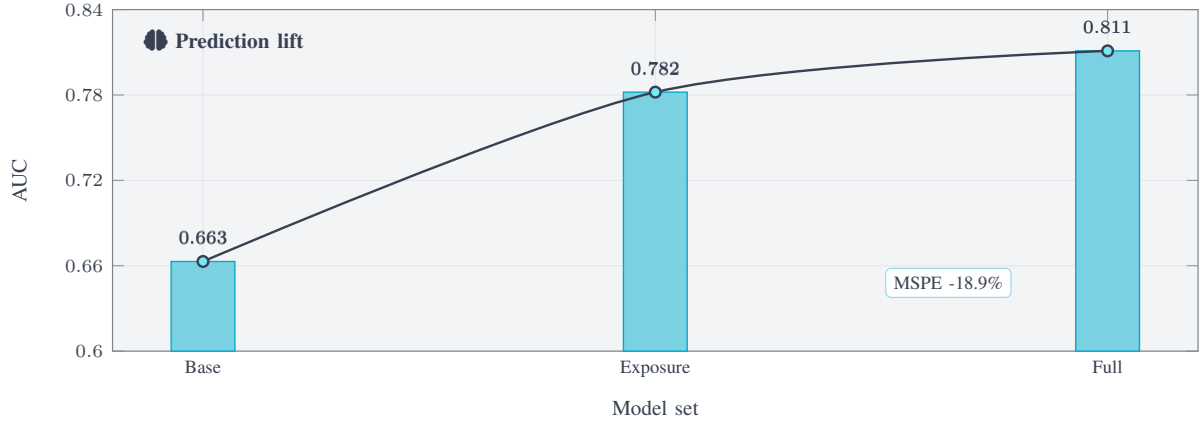


FIGURE 11: Exposure architecture materially improves market-loss prediction.

endorser fame. The evidence suggests that fame is less informative than the structure of the endorsement relationship.

The causal forest estimates show strong heterogeneity. In the highest vulnerability quintile, high sponsorship exposure reduces abnormal returns by 1.86%. These cases combine recent campaign use, high creative dependence, high fit, high audience overlap, and high consensus negativity [33]. In the lowest vulnerability quintile, the estimated effect is minus 0.09% and not statistically significant. The difference between quintiles is large and stable across random forest specifications.

The stakeholder contamination forest produces similar but not identical heterogeneity. The strongest contamination occurs when high audience overlap, high fit, and high controversy moralization coincide. Creative dependence matters less for sentiment than for market reaction. This difference is sensible. Consumers need to recognize the endorser-brand link, but investors additionally price the cost of replacing creative assets. Thus, the market and stakeholder models share some moderators but differ in operational cost sensitivity.

Political polarization events display the most complex heterogeneity. In low-consensus controversies, some sponsors experience negative sentiment among one consumer group and positive sentiment among another. The average effect is small because opposing reactions offset each other. The causal forest identifies brands with broad mainstream positioning as more vulnerable to polarizing endorser behavior than brands with niche or culturally aligned audiences. For mainstream brands, political controversy creates customer-base fragmentation risk. For niche brands, the same controversy may be less damaging or even reinforcing.

Deceptive commercial conduct shows a different pattern. Spillover is severe when the endorser promoted the sponsor's product using performance claims, discount codes, financial claims, health claims, or personal-use testimonials. If the deception is unrelated to the sponsor's product, the spillover is smaller. This finding supports a relevance mechanism. Stakeholders punish the sponsor most when the endorser's credibility was directly used to persuade consumers about the product.

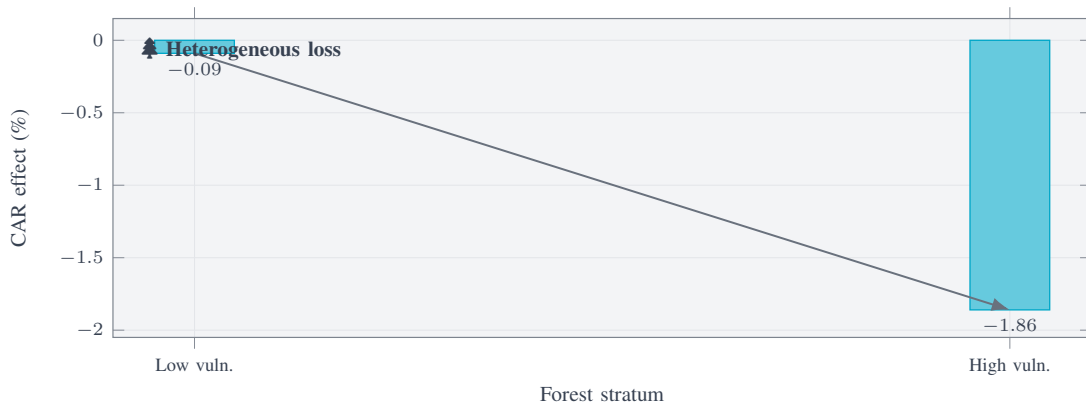


FIGURE 12: Causal-forest strata reveal concentrated downside risk.

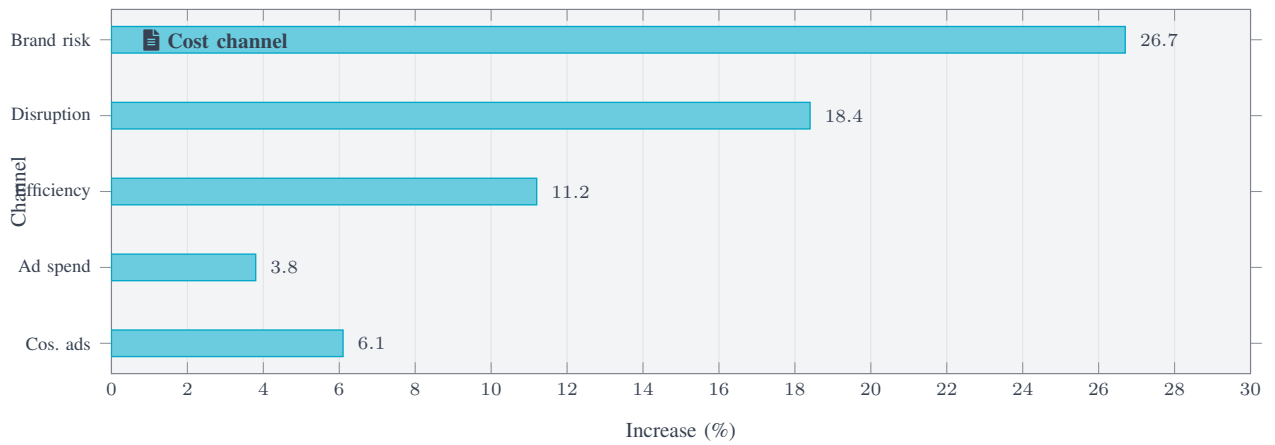


FIGURE 13: Analyst language and advertising spend reflect replacement costs.

Legal allegations have the largest uncertainty effects. Option-implied volatility rises more for legal allegations than for speech controversies with similar sentiment effects. The reason appears to be ambiguity about future information, contract obligations, and consumer response. Legal allegations often unfold over time, and sponsors may delay action until facts become clearer. Option markets price that uncertainty even when immediate sentiment is not yet fully negative.

Response-specific heterogeneity is also clear. Termination is most effective for high-consensus negative events and low creative dependence. It is less effective when creative dependence is high because replacement costs remain. Pause-and-review is most effective when legal facts are uncertain and consensus negativity is moderate. Governance disclosure is most effective after deceptive commercial conduct because it reassures stakeholders that endorsement screening and claim review will change. Replacement campaigns are most effective in seasonal categories with high advertising cadence.

A policy-learning exercise estimates expected recovery under alternative response strategies. For high-consensus legal or discriminatory controversies, predicted brand contamination after six weeks is lowest under termination plus replacement campaign. For uncertain legal allegations with moderate exposure, predicted contamination is lowest under pause plus factual review disclosure. For low-exposure controversies, no response or minimal clarification often avoids sponsor-centered attention. This simulation is descriptive and model-based, not a prescriptive rule, but it illustrates the value of conditioning response on exposure architecture.

The machine-learning analysis also identifies false positives. Some high-exposure sponsors avoid negative reactions because the endorser's fan community defends the sponsor, because the brand's audience is culturally aligned with the endorser, or because the firm had already reduced campaign visibility before the controversy. False negatives often involve hidden contracts or affiliate relationships not captured in public data. These cases indicate that exposure measurement remains imperfect, especially in influencer marketing where relationships can be informal or poorly disclosed.

The strongest conclusion from the machine-learning section is that endorser controversy spillover is predictable only when models understand the endorsement relationship. A model that sees controversy severity but not campaign architecture mistakes public attention for sponsor exposure. A model that sees sponsorship exposure but not consensus negativity mistakes ordinary partnership for reputational risk. The spillover mechanism depends on the intersection of controversy type, stakeholder interpretation, and campaign structure.

VII. Robustness and Additional Evidence

The first robustness test changes the event window. The two-day abnormal return window produces a high-exposure penalty of minus 0.63%. The three-day window produces minus 0.81%. The five-day window produces minus 0.93%. The eleven-day window produces minus 1.02% but includes more confounding news. The growing magnitude in longer windows is consistent with gradual diffusion from social media to business press and investor interpretation.

The second robustness 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 exposure coefficients. The coefficient on sponsorship exposure ranges from minus 0.41% to minus 0.53% per standard deviation. The preferred model produces minus 0.48%. The result is not sensitive to the asset-pricing benchmark.

The third test restricts the sample to high-confidence sponsorships with contract evidence or repeated official brand amplification. The sample falls from 1,138 controversies to 742. The exposure coefficient becomes larger, at minus 0.57% per standard deviation. This pattern is consistent with measurement error in the broader sample, where some weaker relationships dilute the estimated effect.

The fourth test removes endorsers with more than five simultaneous sponsors. Highly sponsored endorsers may spread reputational damage across many firms, making firm-specific exposure harder to identify. Removing them leaves the main coefficient nearly unchanged. When simultaneous sponsors remain in the sample, exposure architecture explains which sponsors are penalized more. The market does not punish all sponsors equally simply because they share an endorser.

The fifth test examines dormant endorsements. Relationships with no observed campaign activity in the six months before controversy produce no significant abnormal returns, even when the formal contract may still exist. This supports the recency component of exposure architecture. Stakehold-

ers react to visible association, not to obscure historical contracts.

The sixth test estimates within-firm models for firms with multiple endorser events. This specification compares different endorser controversies within the same sponsoring firm. The exposure coefficient remains negative. This reduces concern that the result is caused by stable differences between firms that choose endorsers and firms that do not. Within the same firm, highly exposed endorser relationships create larger penalties than low-exposure relationships [34].

The seventh test examines endorser death or illness events as non-culpable shocks. These events create attention and sometimes campaign disruption, but they do not create the same negative sentiment [35]. Market reactions are small and sometimes positive when memorial campaigns preserve brand goodwill. This comparison supports the reputational interpretation. The penalty is not caused solely by unexpected endorser unavailability; it is caused by negative stakeholder evaluation of the association.

The eighth test separates firm response timing from response content. Quick responses are not uniformly beneficial. Fast but vague responses do not improve recovery. Delayed but specific responses are better than immediate generic statements in uncertain legal cases. However, in high-consensus negative events, delayed response worsens both sentiment and implied volatility [36]. Timing and specificity must therefore be modeled jointly. The best response depends on whether facts are uncertain and whether the endorser remains visible in current campaigns.

The ninth test evaluates product involvement. High-involvement products, such as financial apps, health products, and performance equipment, show stronger spillover for deceptive conduct and misinformation. Low-involvement products show stronger spillover for discriminatory speech only when the endorser appears in mass advertising. This suggests that stakeholder concern is category-specific [37]. In high-involvement products, endorser credibility is tied to trust. In low-involvement products, endorser identity and cultural meaning matter more.

The tenth test examines investor base. Firms with higher institutional ownership react more quickly, with less post-event drift. Firms with lower analyst coverage show slower adjustment and larger five-day reactions. Option-market reactions are strongest when option liquidity is high, but the direction remains similar in lower-liquidity firms using realized volatility. Investor sophistication affects the speed of reaction more than the sign of reaction.

The eleventh test uses analyst text as an outcome [38]. High-exposure severe controversies increase analyst mentions of brand risk by 26.7%, campaign disruption by 18.4%,

and advertising efficiency by 11.2%. Analysts mention contract termination costs primarily in high creative-dependence cases. Analyst reports rarely discuss low-exposure influencer controversies, even when social-media attention is high. This validates the market result: professional investors care about controversy when it is connected to brand economics.

The twelfth test analyzes advertising spend after controversy. Firms with high-exposure controversies increase category-adjusted advertising spend by 3.8% over the next quarter, largely due to replacement campaigns and media reallocation. The increase is 6.1% in apparel and cosmetics. This supports the cost channel. A sponsor may suffer not only demand softness but also additional expenditure to rebuild or replace campaign assets.

The thirteenth test examines long-term brand sentiment. For most events, brand contamination returns close to baseline within twelve weeks. Legal allegation events with high creative dependence persist longer. Deceptive commercial conduct tied directly to the sponsor's product also persists. Political polarization decays quickly in aggregate but remains segmented across consumer groups. The long-term sentiment evidence indicates that most endorsement controversies are transitory, but a subset leaves durable residue.

The fourteenth test studies campaign removal from public digital assets. Brands that remove endorser material within five days experience faster reduction in brand contamination when exposure is high. Removal has little effect when exposure is low, because stakeholders were not strongly linking the brand to the endorser. Removal is less effective when screenshots, user reposts, packaging, or retail displays remain visible. Digital removability is therefore only part of the creative-dependence problem.

The fifteenth test examines co-branded products. Co-branded products have the largest and slowest spillovers. A controversy involving a co-branded endorser reduces abnormal returns by 1.62% on average for the sponsoring firm and raises brand contamination by 0.96 standard deviations. Product discontinuation announcements reduce sentiment faster but increase analyst concern about inventory write-downs. This is the clearest case where reputational and operational costs interact [39].

The sixteenth test analyzes affiliate and influencer disclosure quality. Controversies involving creators with weak sponsorship disclosure generate stronger stakeholder backlash when the sponsor had promoted the creator's product claims [40]. This is especially true in cosmetics, health products, and financial applications. Poor disclosure makes the relationship appear less transparent and increases perceptions that the firm benefited from ambiguous persuasion. Sponsorship governance therefore includes disclosure discipline, not only endorser selection.

The seventeenth test uses alternative stakeholder sentiment models. A general sentiment classifier produces similar but noisier results. A controversy-specific classifier produces stronger estimates because it distinguishes ordinary product dissatisfaction from controversy-linked brand contamination. The main findings remain when only posts explicitly naming both the sponsor and endorser are used. That restricted measure is conservative but produces the same category ranking.

The eighteenth test examines international versus domestic endorser controversies. Domestic controversies produce larger immediate market reactions for U.S.-listed firms, but international controversies produce significant reactions when the brand has large exposure in the endorser's home market or when the campaign is global. Global campaigns with uniform creative assets are more vulnerable than locally segmented campaigns. Geographic segmentation can therefore limit spillover.

The nineteenth test assesses legal disclaimers in endorsement contracts. Firms that disclose morality-clause frameworks or endorsement-risk governance in prior filings experience smaller implied-volatility increases after severe controversies. The stock-return moderation is weaker. This suggests that contract governance reduces uncertainty about response options but does not fully prevent reputational harm. Investors value evidence that the firm can act, but consumers still react to the visible association.

The twentieth test evaluates repeated endorser controversies within a firm. Firms with multiple controversy-linked endorsers over a three-year period experience larger penalties for later events [41]. The first event has a high-exposure penalty of minus 0.67%, the second minus 0.98%, and the third minus 1.31%. Repetition appears to change investor interpretation from isolated endorser risk to weak sponsorship governance. This repeated-event pattern is strongest in influencer-heavy categories.

VIII. Implications

The findings indicate that endorsement risk is not reducible to endorser fame or controversy severity. The crucial variable is how the endorser is connected to the brand at the moment of controversy. Recent visibility, high audience overlap, strong symbolic fit, and creative dependence make reputational separation difficult. These attributes are often desirable when the campaign is launched. They increase persuasion, recognition, and identity transfer. Yet they also increase spillover when the endorser becomes controversial. Endorsement value and endorsement risk arise from the same architecture [42].

For managers, the evidence suggests that endorsement decisions should be evaluated as portfolio-risk decisions. A firm that uses many low-dependence creators has different exposure than a firm built around a few central ambassadors. The latter may generate stronger brand meaning, but it also requires deeper screening, contract flexibility, and replacement planning. Screening should not focus only on past legal problems. It should also consider audience polarization, claim credibility, disclosure history, and fit with sensitive product categories [43].

The findings also imply that response strategy should be conditional [44]. There is no single superior response [45]. High-consensus negative events with high exposure require rapid campaign-status clarity. Politically polarized events require more careful audience analysis because termination can become a second signal. Deceptive conduct requires governance disclosure and claim-review correction. High creative dependence requires replacement planning, not merely distancing language. Low-exposure events may be best handled with limited response because excessive sponsor commentary can increase association.

For investor relations, endorser controversy is most important when the endorser is tied to growth, product launch, or category trust. Investors are less concerned with ordinary celebrity noise than with campaign disruption, advertising replacement cost, consumer sentiment, and governance weakness. A firm facing a severe high-exposure controversy can reduce uncertainty by clarifying the campaign's status, contractual flexibility, expected replacement plan, and financial materiality [46]. Ambiguity tends to appear in option-implied volatility even when stock returns are moderate.

For researchers, the evidence shows that reputation spillovers can be measured through relationship architecture. A binary sponsor indicator is too crude. It misses recency, fit, dependence, and audience overlap. The same reasoning applies to other reputational relationships, such as licensing, co-branding, franchise partnerships, expert certification, and creator commerce. Reputation transfer is not simply present or absent. It has structure, intensity, and removability.

For public policy and platform governance, the results highlight disclosure issues in influencer and affiliate marketing. Stakeholder backlash is stronger when prior disclosure was weak and the controversy involves deceptive commercial conduct. This suggests that transparent sponsorship labeling can reduce later reputational confusion. Clear disclosure does not prevent all harm, but it clarifies the relationship and reduces perceptions that the sponsor benefited from hidden persuasion.

The evidence also speaks to brand equity management. A strong brand may absorb minor endorser noise, but it is not fully protected from high-fit controversy. In some settings,

strong brand identity can increase disappointment when the endorser's conduct conflicts with the brand's stated values. Brand strength is therefore not a shield against all symbolic asset failures. It is a resource whose protection requires consistency between partners and positioning.

The study has limitations. Sponsorship contracts are not always public, and some influencer arrangements are intentionally opaque. Audience overlap is estimated from platform data and may contain measurement error. Demand proxies are available only for some product categories. Market reactions are observed at the public parent level even when the endorsement concerns a smaller brand. The estimates may understate effects for private brands and overstate effects for firms whose parent stock is dominated by unrelated segments.

The design is observational. Firms select endorsers strategically, and high-exposure campaigns may differ from low-exposure campaigns in unobserved ways. Firm fixed effects, event controls, placebo tests, and double machine learning reduce but do not remove this concern. The results should be interpreted as adjusted market and stakeholder reactions to public endorser controversy exposure, not as randomized causal effects of endorsement design.

Future research could extend the analysis with confidential contract terms, creative-production budgets, product-level sales, experimental stakeholder responses, and cross-country endorsement norms. Another useful extension would examine insurance, morality clauses, and contract-option design as mechanisms for managing endorser risk. A further extension could study artificial influencers and virtual endorsers, where firms may reduce human controversy risk while creating new authenticity concerns.

IX. Conclusion

This study examined reputation and market reactions to public controversies involving celebrity, athlete, expert, and digital creator endorsers. The evidence shows that controversy spillover is not determined by controversy severity alone. Sponsoring firms experience the largest penalties when the endorser was recent, visible, audience-matched, symbolically aligned with the brand, and difficult to remove from campaign assets. These elements form the sponsorship exposure architecture that governs reputational transmission.

The market results show that high-exposure controversies reduce abnormal returns, raise option-implied volatility, increase trading volume, and affect analyst language. The strongest penalties occur for legal allegations, discriminatory or abusive speech, deceptive commercial conduct, and health or safety misinformation. Political polarization has more variable effects because audience reactions can offset

each other. Professional-performance scandals matter mainly when product claims depend on performance credibility.

Stakeholder outcomes help explain the financial reactions. Brand contamination rises when consumers actively connect the sponsor to the endorser controversy. Search behavior, product-review text, social sentiment, and purchase proxies show that the spillover can move beyond attention into consumer behavior, especially in categories built on identity, trust, or personal aspiration. The demand effects are usually moderate, but they are measurable in high-exposure severe cases.

Firm response matters, but not in a simple way. Specific responses that clarify campaign status, replacement steps, and endorsement governance shorten sentiment recovery [47]. Termination is most useful when the controversy has broad negative consensus. Pause-and-review is more useful when facts are uncertain. Vague concern statements have limited value. Silence can work for low-exposure relationships but is costly when the endorser remains visible in current campaigns.

The paper's central conclusion is that endorsement creates a tradable symbolic asset with embedded downside risk. The same design choices that make an endorser persuasive also make the sponsor vulnerable when the endorser becomes controversial. Reputation management therefore requires measuring not only who the endorser is, but how deeply the endorser is built into the brand's meaning, audience, and creative system. Future empirical work can extend this approach to other human-asset relationships and to emerging forms of creator-led commerce.

REFERENCES

- [1] L. Dubova, L. Slavíková, J. Azevedo, J. Barstad, P. Gatto, J. Lesinski, D. Pettenella, and R. Stokken, *Review of Policy Instruments for Climate-Smart Mountain Forestry*, pp. 477–506. Springer International Publishing, 11 2021.
- [2] N. Yusoff, M. Alias, and N. Ismail, “Drivers of green purchasing behaviour: a systematic review and a research agenda,” *F1000Research*, vol. 12, pp. 1286–1286, 10 2023.
- [3] S. Brammer, H. He, and K. Mellahi, “Corporate social responsibility, employee organizational identification, and creative effort: The moderating impact of corporate ability,” *Group & Organization Management*, vol. 40, pp. 323–352, 12 2014.
- [4] J. Jayabalan, M. Dorasamy, M. Raman, M. Sambasivan, and S. Harun, “Unleashing frugal innovation in private higher education institutions via intellectual capital and information technology capability: a systematic literature review,” *F1000Research*, vol. 10, pp. 1109–1109, 11 2021.
- [5] S. R. Karna and D. K. Gupta, “Fostering egovernment as state social responsibility (ssr): Case study of an australian city council,” *JeDEM - eJournal of eDemocracy and Open Government*, vol. 4, pp. 318–337, 12 2012.
- [6] T. M. Essawy, “Assessing the application of corporate social responsibility in egyptian travel agencies,” *Journal of Association of Arab Universities for Tourism and Hospitality*, vol. 17, pp. 90–109, 12 2019.
- [7] 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.
- [8] G. N. Prabhu and P. S. Aithal, “Corporate social responsibility – a case study on strategies of indian private and public sector banks,” *International Journal of Case Studies in Business, IT, and Education*, vol. 5, pp. 162–183, 10 2021.
- [9] M. T. Bianchi, “Integrated welfare systems and disclosure: Approaching emerging issues,” *Journal of Modern Accounting and Auditing*, vol. 13, 8 2017.
- [10] D. Chen and Z. Peng, “Altruistic motivation or instrumental motivation: The spatial spillover effect of corporate social responsibility on corporate green innovation,” *PloS one*, vol. 18, pp. e0290125–e0290125, 10 2023.
- [11] A. G. SOLIMINI, B. CAPUTO, M. MANICA, A. D. TORRE, and F. Rizzo, “Poster sessions,” *Tropical Medicine & International Health*, vol. 20, pp. 171–441, 9 2015.
- [12] 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.
- [13] S. Looser, G. Sustainability, S. Mohr, and W. Wehrmeyer, “Crises communication vs. united nation sustainable development goals: The 7s paradigm as feasible solution facilitator,” *International Journal of Crisis Communication*, vol. 3, pp. 62–80, 12 2019.
- [14] P. Vang, *Good Guys : A Cultural Semiotic Study of the Print Advertising of the Oil Industry (1900-2000) - Good Guys A Cultural Semiotic Study of the Print Advertising of the Oil Industry (1900-2000)*. Linköping University Electronic Press, 5 2014.
- [15] L. M. A. Aladwey and R. A. Alsudays, “Does the cultural dimension influence the relationship between firm value and board gender diversity in saudi arabia, mediated by esg scoring?,” *Journal of Risk and Financial Management*, vol. 16, pp. 512–512, 12 2023.
- [16] E. Akkartal and G. Y. Aras, “Sustainability in fleet management,” *Journal of Advanced Research in Economics and Administrative Sciences*, vol. 2, pp. 13–39, 9 2021.
- [17] Z. Patora-Wysocka and Łukasz Sułkowski, “Sustainable incremental organizational change—a case of the textile and apparel industry,” *Sustainability*, vol. 11,

- pp. 1102–, 2 2019.
- [18] J. Wang, H.-J. Ban, H.-W. Joung, and H.-S. Kim, “Navigations for hospitality human resource management research: Observing the keywords, factors, topics under the covid-19 pandemic,” *Information*, vol. 13, pp. 126–126, 3 2022.
 - [19] S. H. Kukkuhalli, “Enabling customer 360 view and customer touchpoint tracking across digital and non-digital channels,” *Journal of Marketing & Supply Chain Management*, vol. 1, no. 3, 2022.
 - [20] M. Jankalová and R. Jankal, “The assessment of corporate social responsibility: approaches analysis,” *Entrepreneurship and Sustainability Issues*, vol. 4, pp. 441–459, 6 2017.
 - [21] R. M. Ali*, M. I. Afzal, S. Ahmed, A. S. B. Mahomed, and N. A. Brohi, “The impact of csr practices of mncs on dairy industry in malaysia,” *International Journal of Innovative Technology and Exploring Engineering*, vol. 9, pp. 638–648, 12 2019.
 - [22] A. Barlas, A. Valakosta, C. Katsionis, A. Oikonomou, and V. Brinia, “The effect of corporate social responsibility on customer trust and loyalty,” *Sustainability*, vol. 15, pp. 1036–1036, 1 2023.
 - [23] D. Daugaard and A. Ding, “Global drivers for esg performance: The body of knowledge,” *Sustainability*, vol. 14, pp. 2322–2322, 2 2022.
 - [24] M. M. Thottoli, “The interrelationship of marketing, accounting and auditing with corporate social responsibility,” *PSU Research Review*, vol. 7, pp. 117–136, 5 2021.
 - [25] 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.
 - [26] E. IJBM, “International journal of business and management, vol. 3, no. 10, october 2008,” *International Journal of Business and Management*, vol. 3, 2 2009.
 - [27] G. Zarei and K. Mirzaei, “Impact of perceived social responsibility on consumers’ green perception and green behavior: The moderating role of personality traits,” *Journal of Consumer Sciences*, vol. 7, pp. 168–188, 8 2022.
 - [28] X. Tao, S. Sindakis, C. Chen, P. Theodorou, and S. Showkat, “Validation analysis of charitable organizations and media monitoring using an evolutionary model in china,” *Journal of the knowledge economy*, vol. 15, pp. 1–5570, 4 2023.
 - [29] *The New Microfinance Handbook*. The World Bank, 2 2013.
 - [30] A. Karman and A. Savanevičienė, “Enhancing dynamic capabilities to improve sustainable competitiveness: insights from research on organisations of the baltic region,” *Baltic Journal of Management*, vol. 16, pp. 318–341, 11 2020.
 - [31] Z. A. Mahmoud and H. M. Shamma, “Does corporate reputation affect brand preference? application to the soft drinks industry in egypt,” *African Journal of Marketing Management*, vol. 7, pp. 7–15, 1 2015.
 - [32] Z. Cao, “Brand equity, warranty costs, and firm value,” *International Journal of Research in Marketing*, vol. 39, no. 4, pp. 1166–1185, 2022.
 - [33] P. N. Muriithi, T. Mwanthi, and N. Chiroma, “The triple bottom line and church sustainability in kenya,” *European Journal of Humanities and Social Sciences*, vol. 2, pp. 31–36, 9 2022.
 - [34] S. Mena, M. de Leede, D. Baumann, N. Black, S. Lindeman, and L. McShane, “Advancing the business and human rights agenda: Dialogue, empowerment, and constructive engagement,” *Journal of Business Ethics*, vol. 93, pp. 161–188, 10 2009.
 - [35] J. H. Lee, J. Wood, and J. Kim, “Tracing the trends in sustainability and social media research using topic modeling,” *Sustainability*, vol. 13, pp. 1269–, 1 2021.
 - [36] S. M. Brondoni and L. Bosetti, “Ouverture de ‘integrated csr management’,” *Symphonya. Emerging Issues in Management*, pp. 1–17, 7 2018.
 - [37] W. C., P. G., and L. S., “Organizational culture as a precursor to the practice of corporate social responsibility: A conceptual standpoint,” *British Journal of Management and Marketing Studies*, vol. 4, pp. 62–72, 8 2021.
 - [38] S. Tariq, M. S. Yunis, S. Shoaib, F. Abdullah, and S. W. Khan, “Perceived corporate social responsibility and pro-environmental behaviour: Insights from business schools of peshawar, pakistan,” *Frontiers in psychology*, vol. 13, pp. 948059–948059, 7 2022.
 - [39] L.-H. C. Hsieh and D. S. Noonan, “Strategic behavior in certifying green buildings: An inquiry of the non-building performance value,” *Environmental management*, vol. 60, pp. 231–242, 5 2017.
 - [40] K. Rimita, S. N. Hoon, and R. E. Levasseur, “Leader readiness in a volatile, uncertain, complex, and ambiguous business environment,” *Journal of Social Change*, vol. 12, pp. 2–, 1 2020.
 - [41] J. Jensen, *Whose Rules? Globalizing Governance and the Great Disruption*. Sciyo, 10 2010.
 - [42] D. D. Peña-Miranda, J. Arteaga-Ortiz, and J. Ramón-Cardona, “Determinants of csr application in the hotel industry of the colombian caribbean,” *Sustainability*, vol. 11, pp. 5045–, 9 2019.
 - [43] S. Niu, B. I. Park, and J. S. Jung, “The effects of digital leadership and esg management on organizational innovation and sustainability,” *Sustainability*, vol. 14, pp. 15639–15639, 11 2022.
 - [44] J. Jayabalan, M. Dorasamy, M. Raman, M. Sambasivan, and S. Harun, “Unleashing frugal innovation in private higher education institutions via intellectual capital: a systematic literature review,” *F1000Research*, vol. 10, pp. 1109–, 11 2021.

- [45] W.-E. M. Ibrahim, "Societal responsibility of the foreign direct investment companies in the arabian," (), vol. 39, pp. 227–251, 12 2019.
- [46] K. Anderson, *Australia's Economy in its International Context: The Joseph Fisher Lectures, Volume 2*. University of Adelaide Press, 4 2012.
- [47] V. Mignenan, "Innovation management and overall business performance: Exploratory study in chadian context in the two logones," *International Journal of Business Administration*, vol. 13, pp. 18–18, 2 2022.