Apex G-Score™ Backtest Results: Full Methodology and Evidence (2020–2025)

Yunjung (Michelle) You, Ph.D.
Apex Governance LLC | April 2026

This document is provided for informational and research purposes only. It does not constitute investment advice. Past governance scores do not guarantee future financial performance.

1. Executive Summary

• We scored **2,099 companies** across KOSPI (844) and KOSDAQ (1,255) using only DART regulatory filings—no surveys, no self-assessments, no management interviews.

• Across **nine independent scoring periods** (2020–2025), the top-quintile G-Score firms outperformed the bottom quintile in ROE in **every single period**. Average Q5–Q1 gap: **+12.4 percentage points**.

• Archetype-level analysis shows a **27.5-point cumulative ROE gap** between Celestial and Time Bomb firms over two years (2024–2025).

• Each axis predicts a **distinct financial outcome**: T predicts profitability (r=+0.19, p<0.001), B predicts loss avoidance (r=–0.126, p<0.01), R predicts leverage (r=–0.15, p<0.001). None redundant.

• Kill Switch firms carry **2.7× the debt ratio** of non-KS firms (433% vs. 162%, p=0.03) and exhibit materially higher rates of loss, debt deterioration, and revenue decline in subsequent years.

2. Why a New Governance Measure for the Korean Market

2.1 The Korea Discount: A Structural Problem, Not an Industry Problem

Korea's capital market has achieved remarkable growth over the past several decades. Globally competitive firms have emerged across manufacturing, IT, and biotech, and by measures of market size, liquidity, and corporate earnings, Korea has long surpassed the emerging-market category. Yet Korean equities have persistently traded at lower valuations than global peers with comparable fundamentals—a phenomenon known as the "Korea Discount." Empirical analysis across 28 countries confirms that Korean stocks carry significantly lower price-earnings ratios than their international counterparts, even after controlling for industry and firm characteristics (Ducret & Isakov, 2020).

This discount is often attributed to geopolitical risk, macroeconomic factors, or declining industrial competitiveness. However, these explanations are insufficient to account for the persistence and scale of the undervaluation relative to Korea's economic trajectory. The core issue, from the perspective of global capital allocators, is not whether Korean firms can generate returns—but whether those returns will accrue to shareholders in a predictable and equitable manner. The Korea Discount reflects not a failure of Korean industry, but a structural uncertainty premium imposed by investors who lack confidence in the governance mechanisms that determine how corporate value is distributed (Black, Jang & Kim, 2006; Joh, 2003).

This uncertainty is not driven by transient external shocks. It is the accumulated result of decades of opacity in corporate decision-making structures, information asymmetry between management and shareholders, and

inadequate investor protection mechanisms. As a consequence, Korea's equity market bears a persistent risk premium that its economic fundamentals do not warrant.

2.2 The Structural Mechanics of Korean Governance Risk

Korean governance risk cannot be characterized simply as "lack of transparency." The problem is more specific and more structural: information exists, but shareholders cannot systematically track, connect, or verify the patterns within it. Audit reports and regulatory filings are publicly available. Yet critical decisions and related-party transactions are disclosed as isolated events—summaries of outcomes, stripped of context, repetition, and structural significance.

Information asymmetry as structural opacity, not data absence. When the same type of related-party transaction recurs across multiple fiscal years, each instance is disclosed separately. The cumulative effect—a systematic pattern of capital transfer to specific stakeholders—is never presented as such. Individual transactions appear "one-off"; in aggregate, they constitute a structural flow. Information is disclosed, but not in a form that enables shareholders to make informed judgments. This is the gap that the academic literature on "tunneling"—the transfer of assets out of firms for the benefit of controlling shareholders—has documented extensively (Johnson, La Porta, Lopez-de-Silanes & Shleifer, 2000).

Legal form masking concentrated power. Much of Korean governance risk arises not from illegal acts but from legally permissible structures that concentrate decision-making power. Boards and audit committees exist in form, but substantive authority is held by a small number of stakeholders. External checks operate nominally. Within this framework, transactions favorable to controlling shareholders or related parties can persist for years with full procedural compliance. Value destruction occurs gradually and cumulatively; by the time illegality is identified ex post, economic losses have already materialized. Research on Korean chaebols has documented how mergers and related-party transactions within business groups systematically benefit controlling shareholders at the expense of minority investors (Bae, Kang & Kim, 2002).

Asset diversion as pattern, not event. Shareholder value erosion rarely occurs through a single catastrophic act. It emerges through repeated related-party transactions with asymmetric terms, minimal or formulaic disclosure, and long-term entrenchment of unbalanced governance structures. Each transaction, viewed in isolation, may appear unobjectionable. Viewed over time, the cumulative capital flow from the firm to specific stakeholders becomes clear—yet no existing evaluation framework is designed to detect this temporal pattern.

Shareholders as passive recipients. In this structure, shareholders cannot access decision-making processes, cannot contextualize capital movements in advance, and are left to absorb outcomes after the fact. This is not merely an information problem—it is a structural limitation of shareholder authority embedded in the governance architecture itself.

2.3 Why Existing Evaluation Frameworks Fail to Capture This Risk

Korea's governance problems have been addressed through ESG ratings and governance evaluation systems for decades. Many firms receive acceptable or high governance scores. Yet shareholder value destruction and structural risk events have continued to recur. This disconnect is not accidental—it reflects fundamental architectural limitations in existing frameworks. A comprehensive survey of 108 academic articles published between 2011 and 2020 on Korean corporate governance confirms that these challenges remain empirically unresolved (Kim, Kim & Park, 2021).

- **Self-reported inputs.** Most frameworks rely on company-submitted surveys, reports, and policy statements. The entity being evaluated controls the inputs to its own evaluation. Research on investor protection has established that governance quality depends on enforcement and external verification, not self-declaration (La Porta, Lopez-de-Silanes, Shleifer & Vishny, 2000). These frameworks ask "does the institution exist?" but not "does it function?"

- **Formal compliance over functional reality.** Existing ratings credit the presence of governance mechanisms (audit committees, independent directors) without assessing whether they operate. A board that has never recorded a dissenting vote in five years receives the same score as one with active deliberation—if both meet formal criteria. Research using Korean panel data has shown that governance affects firm value specifically through board structure's functional operation, not through procedural box-checking (Black, Kim, Jang & Park, 2015).
- **Single-score aggregation.** Collapsing transparency, power balance, and conflict-of-interest risk into one number creates false equivalences. The seminal G-Index (Gompers, Ishii & Metrick, 2003) demonstrated that a governance index can predict equity returns, but its single-score design obscures which dimension drives the result. A company with T=95 and B=30 receives the same composite as one with 62 on both—yet these are radically different risk profiles.
- **Inability to detect cumulative patterns.** Governance risk accumulates over time through repeated transactions, entrenched structures, and persistent asymmetries. Existing frameworks evaluate point-in-time snapshots—a specific year's institutional requirements—and cannot recognize that the same type of value-extractive transaction has recurred across multiple periods.
- **Investor perspective absent.** Existing ESG and governance ratings were developed primarily for regulatory compliance and corporate benchmarking. They ask whether a firm is well-managed by institutional standards—not whether its governance structure protects outside capital over time. The question investors need answered is not "is this company compliant?" but "will this structure protect my capital?"

The cumulative effect of these limitations is a trust deficit: ESG scores are high, but shareholder value is destroyed; governance ratings exist, but risks are not filtered in advance; evaluation results show little correlation with investment outcomes. This perception is not mere skepticism—it translates directly into an elevated risk premium across the Korean equity market.

2.4 The Gap We Address

Global data providers (MSCI, ISS, Bloomberg) apply universal frameworks that are not calibrated for Korea's specific institutional features: chaebol cross-ownership structures, the prevalence of related-party lending, convertible bond dilution mechanics, and the gap between de jure and de facto board independence. KCGS, the domestic evaluator, covers approximately 1,000 firms and does not provide data in API-accessible formats for systematic integration into investment processes.

What is needed is not another governance opinion but a quantitative, verifiable, data-driven measurement system: one that enables cross-firm comparison on a consistent basis, tracks changes over time using the same methodology, and identifies cumulative structural patterns rather than point-in-time snapshots. Governance must transition from the domain of subjective judgment to a measurable investment variable. Risk that is not measured cannot be managed; evaluation that cannot be verified cannot serve as an investment criterion.

The Apex G-Score was designed to fill this gap: a Korea-specific governance scoring system built entirely on publicly verifiable regulatory filings, covering all listed companies, decomposing governance into three independent risk dimensions (Transparency, Balance, Risk), and delivering data in formats that institutional investors can integrate directly into their workflows.

3. Universe and Data

3.1 Scoring Universe

The Apex G-Score is computed for all companies listed on the Korea Exchange (KRX). Our scoring universe comprises **844 KOSPI-listed firms** scored across six fiscal years (2020–2025) and **1,255 KOSDAQ-listed firms** scored for 2023–2025, totaling **2,099 unique companies**.

Exhibit 1: Scoring Universe

Market	Companies Scored	Scoring Period
KOSPI	**844**	**2020–2025 (6 years)**
KOSDAQ	**1,255**	**2023–2025 (3 years)**
Total	**2,099**	

3.2 Data Sources and Integrity

All input data is sourced exclusively from the DART electronic disclosure system operated by Korea's Financial Supervisory Service. The data pipeline ingests annual business reports, corporate governance reports, external audit opinions, shareholder meeting records, and real-time event disclosures. No surveys, self-assessments, or management interviews are used at any stage.

This design choice is deliberate. By restricting inputs to regulatory filings—documents that carry legal liability for misrepresentation—the model eliminates the self-reporting bias inherent in survey-based governance ratings. Every data point used in the G-Score can be independently verified by any market participant with access to DART.

3.3 Backtest Sample

For backtest validation (KOSPI only), we require both governance scores and subsequent-year financial data. The available backtest sample ranges from 489 to 723 firms per period, depending on data availability.

Exhibit 2: Backtest Sample by Scoring Year

Scoring Year	n (with financial data)
2020	489–495
2021	542–563
2022	559–616
2023	714–731
2024	722–723

213 KOSPI companies were excluded from backtesting due to insufficient board-structure disclosure. All excluded firms were small-cap and received D-grade scores. **No large-cap firm was excluded.** This exclusion biases our results conservatively: the excluded firms would likely widen the governance–performance gap if included, given their uniformly low governance quality.

4. Design Philosophy

Before describing the scoring methodology, we set out the design principles that shaped it. These principles are not cosmetic—they are structural constraints that determine what the model can and cannot do, and why.

4.1 Three Axes, Not One

Corporate governance is not a single dimension. As discussed in Section 2, single-index approaches conflate heterogeneous risk drivers into one number. A company that discloses everything but allows minority shareholders no influence has a different risk profile than one that discloses little but distributes power broadly. Collapsing these into a single score erases the diagnostic information that matters most.

The TBR framework (Transparency, Balance, Risk) decomposes governance into three independent axes, each measuring a distinct risk driver. This is not a taxonomic preference—it is an empirical finding. As shown in Section 9, each axis predicts a different financial outcome: T predicts profitability, B predicts loss avoidance, R predicts leverage. If the axes were redundant, the model would be unnecessarily complex. They are not.

4.2 Regulatory Filings as the Sole Input: Eliminating Self-Reporting Bias by Design

The decision to build the G-Score exclusively on DART regulatory filings is not a data sourcing convenience—it is a structural design choice that addresses one of the core failures identified in Section 2. Existing governance frameworks rely on company-submitted surveys, sustainability reports, and management self-assessments. This creates a fundamental architecture problem: the entity being evaluated controls the quality and completeness of the evaluation inputs.

DART filings—annual business reports, audit opinions, governance reports, shareholder meeting records—carry legal liability for misrepresentation. A company can frame a sustainability report however it wishes, but it cannot fabricate an audit opinion or misstate board attendance records in a regulatory filing without legal consequence. By restricting inputs to documents that carry this legal weight, the G-Score eliminates self-reporting bias at the architectural level, not through post-hoc verification.

This design also ensures full verifiability. Every data point used in the G-Score can be independently retrieved by any market participant with access to DART—which is free and publicly available. No proprietary data sources, no gated databases, no information that requires the firm's cooperation to obtain. The model's inputs are as transparent as any governance input can be; what is proprietary is how those inputs are weighted, scored, and combined.

4.3 Absolute Scoring (M0)

The G-Score uses absolute evaluation (M0): a firm's score is determined entirely by its own governance characteristics, not by its relative position within a peer group. This means scores are comparable across time and across sectors without rebasing.

The rationale is straightforward. If a firm's governance worsens but its peers worsen faster, a relative model would show improvement. An absolute model shows deterioration. For investors managing downside risk, the absolute reading is the correct one.

4.4 Annual Scoring with Real-Time Exceptions

The G-Score is computed annually based on year-end filings (typically available in March–April). This reflects the disclosure cadence of the Korean market: most governance-relevant information is updated once per year in the annual business report.

However, a subset of indicators—those related to catastrophic events (embezzlement disclosures, convertible bond issuances, treasury share dispositions)—are monitored in real time via DART's event disclosure feed. When triggered, these can activate the Kill Switch mechanism between annual scoring cycles. This dual cadence mirrors the structure of credit ratings: periodic reviews plus event-driven watch-list actions.

4.5 Verifiability Over Proprietary Formulas

A tension exists in any scoring system between transparency and proprietary value. We resolve it as follows:

What we disclose: the three-axis structure, axis weights (T=30%, B=40%, R=30%), the number of indicators per axis, the types of governance features each axis captures, the archetype taxonomy, and—most importantly—the full backtest results with sample sizes, effect magnitudes, and statistical significance levels. Any researcher can evaluate whether our claimed results are plausible and internally consistent.

What we do not disclose: the specific indicator definitions, item-level scoring rules, point allocations within each axis, the archetype classification algorithm, and Kill Switch trigger conditions. These constitute the intellectual property of the model and are protected as trade secrets.

This is the standard approach in commercial research. MSCI does not publish its ESG scoring formulas; S&P does not disclose its credit rating algorithms. What distinguishes a credible proprietary model from an opaque one is not formula disclosure but **result verifiability**: can an external party evaluate whether the outputs are statistically meaningful, temporally stable, and economically interpretable? The data presented in Sections 6–10 of this note is designed to enable exactly that evaluation.

4.6 Non-Compensatory Architecture

While the composite G-Score is a weighted average of T, B, and R, the model incorporates a non-compensatory override: the Kill Switch. This binary mechanism ensures that catastrophic governance failures—embezzlement, audit opinion refusal, systematic asset diversion—cannot be offset by high scores on other axes. A firm with T=95, B=90, and a Kill Switch trigger is not “mostly good”—it is non-evaluable. This reflects the empirical reality that governance catastrophes are discontinuous events, not gradual declines that a weighted average can capture.

---

5. Scoring Methodology

The Apex G-Score evaluates corporate governance along three independent axes, each scored on a 0–100 absolute scale. The composite score is computed as:

G-Score = T × 0.30 + B × 0.40 + R × 0.30

5.1 Transparency (T) — Weight: 30%

The T-axis measures the timeliness, consistency, and predictability of a company’s information disclosure. It captures not whether a firm discloses—all listed companies are legally required to—but *how* it discloses. Representative indicators include the lead time given to shareholders before critical decisions and the predictability of dividend policies, among others across a total of five indicators.

In the Korean market context, transparency risk is not primarily about data absence. DART provides extensive public filings. The risk lies in structural opacity: information is fragmented across filings in ways that obscure cumulative patterns. A series of individually unremarkable related-party transactions, each disclosed in a separate filing, may constitute a systematic pattern of asset diversion that no single filing reveals. The T-axis is designed to capture the *structural accessibility* of information, not merely its existence.

5.2 Balance of Power (B) — Weight: 40%

The B-axis evaluates whether governance mechanisms that check management power are functioning in practice, not merely present on paper. Representative indicators include shareholder participation mechanisms and audit

committee independence, among others across a total of six indicators.

Balance carries the highest weight (40%) because Korea's governance failures most frequently manifest as power concentration. The structural pattern is well-documented: boards that never record a dissenting vote, audit committees staffed by affiliates of the controlling shareholder, nomination processes that produce predetermined outcomes. These are not violations of law—they are legal structures that concentrate decision-making power in ways that make minority shareholders irrelevant.

The B-axis weight was empirically validated through the rebalancing exercise described in Section 12: shifting weight toward indicators of substantive oversight (and away from formal indicators) significantly improved the axis's predictive power for loss incidence.

5.3 Decision-Making Risk (R) — Weight: 30%

The R-axis tracks structural exposure to conflicts of interest and capital leakage—what the academic literature terms "tunneling" (Johnson, La Porta, Lopez-de-Silanes & Shleifer, 2000). Representative indicators include related-party transaction intensity and convertible instrument dilution risk, among others across a total of five indicators.

These risks are latent under normal conditions but catastrophic when triggered. A firm that has been quietly building related-party lending exposure for three years may show no distress in current financials. The R-axis captures the *structural preconditions* for value destruction—the loaded gun, not the shot.

5.4 Archetype Classification

Beyond the composite score and letter grade (S/A/B/C/D), each firm is classified into one of five governance archetypes based on the interaction of its axis-level scores. The classification formula is proprietary, but the typology is designed to capture qualitatively distinct governance profiles that a single composite score would obscure:

- **Celestial** — Consistently high scores across all three axes. Genuinely well-governed. Financial analysis can proceed with standard assumptions about management alignment.
- **Poison Apple** — High transparency, weak balance and risk controls. The most dangerous category for investors relying on disclosure quality as a governance proxy. These firms look healthy by conventional screens; their structural fragility is invisible to single-score ratings.
- **Hidden Gem** — Low external visibility but sound internal structure. Governance quality is underpriced by the market, often due to weak IR rather than weak governance.
- **Chameleon** — Mixed signals across axes; direction not yet fixed. Requires ongoing monitoring rather than static classification.
- **Time Bomb** — Governance mechanisms that enable rapid concentration of control or destruction of minority value. Kill Switch may apply.

5.5 Kill Switch

The G-Score incorporates a binary override mechanism (Kill Switch) that identifies firms exhibiting governance characteristics associated with catastrophic risk. When triggered, the Kill Switch supersedes the composite score, classifying the firm as non-evaluable under standard scoring. The specific trigger conditions are proprietary. The Kill Switch is not a score adjustment—it is a binary gate that answers a single question: *Is it possible to evaluate this firm through a scoring framework at all?*

6. Grade Distribution

Exhibit 3: Grade Distribution, KOSPI 2025 G-Score

Grade	Score Range	n	% of Universe
S (≥90)	≥90	0	0.0%
A (≥80)	80–89	23	2.7%
B (≥70)	70–79	137	16.2%
C (≥60)	60–69	228	27.0%
D (<60)	<60	456	54.0%

More than half of KOSPI firms score below 60 (D-grade). No firm achieves S-grade (≥90). This concentration at the bottom of the distribution is itself a finding: it quantifies the scale of governance deficits across the Korean market and explains why the Korea Discount persists.

The quintile spread is monotonically increasing across all tested periods. In the most recent full cross-section (2024→2025, n=723), the Q5–Q1 ROE gap is **+9.4 percentage points**.

7. Period-by-Period Cross-Validation

To test robustness, we repeat the quintile analysis across nine independent scoring–outcome periods, spanning pre-COVID, pandemic, rate-hiking, and post-normalization regimes. Each period uses that year's filings for scoring and validates against subsequent financial performance.

Exhibit 4: Nine-Period Cross-Validation (Q5–Q1 ROE Gap)

Period	n	Q5–Q1 Gap	Market Regime
2020 → 2021	489	+14.1%p	COVID recovery
2020 → 2022	495	+8.5%p	COVID recovery
2021 → 2022	563	+17.6%p	Post-vaccine
2021 → 2023	542	+10.9%p	
2022 → 2023	616	+17.0%p	Rate-hiking cycle
2022 → 2024	613	+12.2%p	
2023 → 2024	714	+10.1%p	Normalization
2023 → 2025	714	+12.0%p	
2024 → 2025	723	+9.4%p	Most recent
Average		**+12.4%p**	
Range		**+8.5 ~ +17.6%p**	
Positive periods		**9 / 9 (100%)**	

The Q5–Q1 gap is **positive in all nine periods**. The signal shows no meaningful degradation across market regimes. The average gap during pre/during-COVID periods (2020–2021 base, 4 periods) was +12.8%p; the post-COVID average (2022–2024 base, 5 periods) was +12.1%p—a difference of 0.7%p, well within sampling variability. This temporal stability is critical: a governance signal that works only in specific market conditions has limited investment utility.

8. Archetype-Level Performance

Exhibit 5: Archetype Distribution and Financial Profile (KOSPI, 2025 G-Score)

Archetype	n	% of Universe	2024 ROE	2-Year Cum. ROE	OPM	Debt Ratio	Loss Rate
Celestial	95	11.3%	+7.7%	+15.1%	+6.6%	123%	7.4%
Poison Apple	28	3.3%	+7.5%	+13.5%	+5.6%	92%	—
Hidden Gem	172	20.4%	+4.5%	+10.2%	—	—	—
Chameleon	**494**	**58.5%**	**+2.4%**	**+4.6%**	**–1.6%**	**108%**	**—**
Time Bomb	**55**	**6.5%**	**–5.6%**	**–12.4%**	**–9.6%**	**155%**	**32.7%**
C–TB Gap			**+13.3%p**	**+27.5%p**	**+16.2%p**		

The Celestial–Time Bomb gap of **27.5 percentage points** in two-year cumulative ROE is the model's sharpest signal. Time Bomb firms exhibit a loss rate of 32.7%—**4.4×** the Celestial rate (7.4%). Revenue growth diverges significantly: Celestial firms grew at +8.7% versus –5.0% for Time Bombs ($p<0.001$).

The Poison Apple Problem

Poison Apple firms merit particular attention because they represent the specific failure mode that the three-axis architecture was designed to detect. Their single-year ROE (+7.5%) nearly equals Celestials (+7.7%), and their debt ratio (92%) is the lowest of any archetype. By any conventional financial screen, these firms appear healthy.

However, their governance structure—high transparency masking concentrated power and elevated conflict-of-interest risk—creates a fragility not visible in current-period financials. This is precisely the failure mode that single-score governance ratings cannot detect: a firm that tells you everything but lets you decide nothing. The gap between what these firms disclose and what they allow is where risk concentrates.

9. Axis-Level Predictive Power

A key validation criterion for the TBR framework is that each axis should predict a *distinct* financial outcome. If all three axes predicted the same thing, the model would be redundant—three dimensions measuring one signal. Exhibit 6 presents bivariate correlations between axis scores and financial metrics (KOSPI, n=630, 2024–2025 average).

Exhibit 6: Axis-Level Correlations with Financial Metrics

Axis	vs ROE	vs ROA	vs OPM	vs Debt Ratio	vs Loss Rate
T (Transparency)	r=+0.19***	r=+0.29***	r=+0.18***	n.s.	r=–0.31***

Axis	vs ROE	vs ROA	vs OPM	vs Debt Ratio	vs Loss Rate
B (Balance)	n.s.	n.s.	n.s.	n.s.	r=–0.126**
R (Risk)	r=+0.10**	n.s.	n.s.	r=–0.15***	n.s.
Total G-Score	r=+0.149**	—	—	—	—

Significance: * p<0.05, ** p<0.01, *** p<0.001. n.s. = not significant.

Exhibit 7: Axis Role Summary

Axis	Predicts	Key Finding
T (Transparency)	Profitability	Strongest single predictor of ROE. Disclosure discipline reflects management discipline.
B (Balance)	Loss prevention	Does not predict earnings levels—but predicts whether a firm falls into losses. A structural safeguard.
R (Risk)	Leverage risk	Lower conflict-of-interest exposure = lower debt. Fewer extraction channels, less aggressive borrowing.

Three axes, three jobs. None redundant. This is by design. The TBR framework captures profitability risk (T), survival risk (B), and leverage risk (R) through distinct, empirically validated channels. Prior research has established that governance affects firm value through multiple channels including reduced self-dealing and improved operational performance (Black, Kim, Jang & Park, 2015). A model that collapsed these into one dimension would lose the diagnostic information that enables archetype classification—and would miss the Poison Apple problem entirely.

10. Kill Switch Analysis

The Kill Switch identifies firms whose governance structures enable rapid concentration of control or destruction of minority value.

Exhibit 8: Kill Switch vs. Non-Kill Switch Firms (KOSPI, 2025 G-Score)

Metric	KS ON (n=17)	KS OFF (n=585)	Ratio / Gap	p-value
Debt Ratio	433%	162%	2.7×	p = 0.03
Revenue Growth	–4.1%	+8.6%	–12.7%p	—
ROE	–3.0%	+2.4%	–5.4%p	p = 0.39 (n.s.)

Exhibit 9: Kill Switch Predictive Power (2023 G-Score → 2024–2025 Outcomes)

Metric	KS ON	KS OFF	Gap
Loss rate	33%	18%	+15.7%p
Debt deterioration	67%	50%	+16.3%p
Revenue decline	50%	38%	+12.1%p

The ROE difference between KS ON and KS OFF groups is not statistically significant (p=0.39), likely due to the small sample (n=17). However, the debt ratio difference is significant (p=0.03), and the forward-looking metrics in Exhibit 9 show that Kill Switch firms are materially more likely to experience losses, debt deterioration, and revenue decline in subsequent years.

The interpretation is directional: **governance data captures these risks before the balance sheet makes them obvious.** Kill Switch firms do not look distressed in their current financials. Their debt ratios are high but not necessarily at crisis levels; their revenues are declining but not collapsing. The governance signal leads the financial signal.

11. KOSPI vs. KOSDAQ: Structural Governance Divergence

Exhibit 10: Axis-Level Score Comparison

Axis	KOSPI	KOSDAQ	Gap
T (Transparency)	53.6	63.1	+9.4
B (Balance)	**46.2**	**33.6**	**–12.6**
R (Risk)	**78.4**	**69.7**	**–8.7**
Total G-Score	**58.1**	**53.9**	**–4.2**

The pattern is counterintuitive: KOSDAQ firms score *higher* on Transparency (+9.4) but *lower* on Balance (–12.6) and Risk (–8.7). Smaller firms have simpler disclosure requirements (fewer filings to miss) while lacking governance infrastructure—independent boards, functioning audit committees—that larger KOSPI firms are more likely to have.

Exhibit 11: Archetype Distribution by Market

Archetype	KOSPI	KOSDAQ
Celestial	11.3%	0.0%
Poison Apple	3.3%	20.7%
Hidden Gem	20.4%	0.1%
Chameleon	58.5%	70.1%
Time Bomb	6.5%	9.1%

KOSDAQ has **zero Celestial firms** and **20.7% Poison Apples**. This concentration represents a systematic risk for investors who equate disclosure quality with governance quality in the KOSDAQ market.

Exhibit 12: Archetype ROE by Market (2024)

Archetype	KOSPI ROE	KOSDAQ ROE	Gap
Poison Apple	+7.5%	+4.0%	–3.5%p
Chameleon	+2.3%	–9.3%	–11.6%p
Time Bomb	–5.6%	–14.1%	–8.5%p

Within comparable archetypes, KOSDAQ firms underperform KOSPI firms across the board. The Chameleon gap (–11.6%p) suggests that governance weaknesses are more consequential in the KOSDAQ environment, where

external monitoring and institutional investor oversight are thinner.

12. Model Calibration: B-Axis Weight Rebalancing

During model development, we tested multiple weight configurations for the six indicators within the B-axis. The original specification distributed weights relatively evenly. Empirical analysis revealed that shifting weight toward indicators of *substantive* oversight and away from formal indicators significantly improved predictive power.

Exhibit 13: Impact of B-Axis Weight Rebalancing

Metric	Before Rebalancing	After Rebalancing
B-axis vs. Loss Rate	r = –0.070 (n.s.)	r = –0.126 (p < 0.01)
Total G-Score vs. ROE	r = +0.130 (p < 0.01)	r = +0.149 (p < 0.01)
Celestial–Time Bomb ROE Gap	+13.3%p	+20.7%p

Three improvements occurred simultaneously: (1) the B-axis became a statistically significant predictor of loss incidence, (2) the composite G-Score's correlation with ROE increased, and (3) the archetype-level ROE spread widened from +13.3%p to +20.7%p.

This result validates the core design thesis: **in the Korean market, it is the substance of governance—not its formal existence—that predicts financial outcomes.** Indicators that measure whether oversight mechanisms actively constrain management behavior showed stronger empirical associations with performance than indicators that measure whether governance structures exist on paper. The specific indicators that gained and lost weight are proprietary, but the direction is clear: functional checks matter more than formal ones.

13. Limitations

We present these limitations not as caveats to be discounted but as boundaries of the model's current validity. Users should calibrate their reliance on G-Score data accordingly.

- **Correlation, not causation.** The documented associations between G-Score and financial outcomes do not establish causal relationships. Good governance may be a consequence rather than a cause of financial health, or both may be driven by unobserved factors such as management quality.
- **Survivorship bias.** Our backtest universe includes only firms that remained listed throughout the scoring and outcome periods. Firms that delisted due to governance failures are excluded, which may understate the true governance–performance relationship.
- **Sample exclusion bias.** 213 KOSPI firms were excluded from backtesting due to insufficient board-structure disclosure. While all were small-cap and D-grade, their exclusion limits generalizability to the smallest segment of the market.
- **KOSDAQ data limitation.** KOSDAQ firms have been scored for only three years (2023–2025), insufficient for robust multi-period backtesting. KOSDAQ results should be treated as preliminary.
- **Static annual scoring.** The G-Score is computed annually based on year-end filings and does not capture intra-year governance changes. Kill Switch events triggered by real-time disclosures are an exception but cover only a subset of governance dynamics.

- **Small Kill Switch sample.** The Kill Switch analysis is based on n=17 firms, which limits statistical power. The debt ratio finding (p=0.03) is significant, but the ROE comparison is not.
- **Single-market scope.** The model is calibrated for the Korean market's institutional environment (chaebol structures, related-party transaction norms, DART disclosure regime). Transferability to other markets has not been tested.

14. What This Note Does and Does Not Claim

This note presents evidence that the Apex G-Score—a governance scoring system built on public regulatory filings—shows a statistically significant, temporally stable, and economically meaningful association with subsequent financial performance across Korean listed companies.

We claim:

- The Q5–Q1 ROE spread is positive in all nine tested periods, averaging +12.4%p.
- Each axis predicts a distinct financial outcome, validating the three-dimensional architecture.
- The Kill Switch mechanism identifies firms with materially higher subsequent risk of loss and debt deterioration.
- The Poison Apple archetype isolates a category of governance risk invisible to single-score ratings.

We do not claim:

- That high G-Score firms will outperform. Governance quality is one input to investment decisions, not the decision itself.
- That the G-Score captures all governance risk. Aspects of governance not reflected in DART filings—board dynamics, management culture, informal power structures—are outside the model's scope.
- That the historical backtest results will persist. Market regimes, regulatory environments, and corporate behavior change.

The G-Score is a pre-analysis filter, not a trading signal. Its value lies in helping investors decide where to allocate analytical resources—and where not to.

15. Suggested Citation

You, Y. (2026). Apex G-Score Backtest Results: Full Methodology and Evidence (2020–2025). *Apex Governance LLC, Research Note #1.*

Appendix A: Executive Compensation by Archetype

Executive pay ratios (CEO total compensation / average employee compensation) are not used as inputs to the G-Score model. They are presented here as an independent reference variable that provides additional context on how governance structures relate to compensation practices.

Exhibit A-1: Pay Ratio Distribution (KOSPI, n=830)

Metric	Value
Mean	4.2×
Median	3.0×
P25	1.9×
P75	4.7×

Exhibit A-2: Pay Ratio by Archetype

Archetype	Mean Pay Ratio	Median Pay Ratio
Poison Apple	7.5×	6.3×
Celestial	5.3×	3.6×
Chameleon	4.1×	2.9×
Hidden Gem	3.5×	2.8×
Time Bomb	3.4×	2.2×

The most notable finding is the position of Poison Apple firms at the top of the pay ratio distribution. Their mean pay ratio (7.5×) exceeds even Celestials (5.3×), while their median (6.3×) is the highest of any archetype by a wide margin. This is consistent with the Poison Apple profile: firms that maintain high disclosure standards while concentrating power—and economic benefit—in management. The compensation data provides an independent signal that corroborates the structural risk identified by the TBR axis scores.

Time Bomb firms, by contrast, have the lowest pay ratios (mean 3.4×, median 2.2×). This likely reflects depressed earnings and smaller firm size rather than egalitarian compensation practices. Low pay ratios in distressed firms should not be interpreted as governance strength.

Appendix B: KOSDAQ Detailed Grade Distribution

Exhibit B-1: Grade Distribution Comparison (KOSPI vs. KOSDAQ)

Grade	KOSPI	KOSDAQ
A (≥80)	2.7%	0.0%
B (70–79)	16.2%	0.6%
C (60–69)	27.0%	19.0%
D (<60)	54.0%	80.5%

The grade distribution divergence between markets is stark. On KOSDAQ, **80.5% of firms receive D-grade** (below 60), compared to 54.0% on KOSPI. No KOSDAQ firm reaches A-grade, and only 0.6% achieve B-grade. This near-complete absence of governance quality in the upper grades suggests that the KOSDAQ market's governance infrastructure is structurally underdeveloped, not merely lagging.

For investors allocating to KOSDAQ, this distribution implies that governance screening is not a refinement of analysis but a prerequisite: the base rate of governance risk is so high that deploying capital without governance due diligence is equivalent to accepting unscreened exposure to the D-grade majority.

References

Bae, K.-H., Kang, J.-K., & Kim, J.-M. (2002). Tunneling or Value Added? Evidence from Mergers by Korean Business Groups. *Journal of Finance*, 57(6), 2695–2740.

Black, B., Jang, H., & Kim, W. (2006). Does Corporate Governance Predict Firms' Market Values? Evidence from Korea. *Journal of Law, Economics, and Organization*, 22(2), 366–413.

Black, B., Kim, W., Jang, H., & Park, K.S. (2015). How Corporate Governance Affects Firm Value: Evidence on Channels from Korea. *Journal of Banking and Finance*, 51, 131–150.

Ducret, R. & Isakov, D. (2020). The Korea Discount and Chaebols. *Pacific-Basin Finance Journal*, 63, 101396.

Gompers, P., Ishii, J., & Metrick, A. (2003). Corporate Governance and Equity Prices. *Quarterly Journal of Economics*, 118(1), 107–155.

Joh, S.W. (2003). Corporate Governance and Firm Profitability: Evidence from Korea before the Economic Crisis. *Journal of Financial Economics*, 68(2), 287–322.

Johnson, S., La Porta, R., Lopez-de-Silanes, F., & Shleifer, A. (2000). Tunneling. *American Economic Review*, 90(2), 22–27.

Kim, W., Kim, W., & Park, K.S. (2021). A Survey of Research on the Corporate Governance of Korean Firms. *ECGI Finance Working Paper No. 804/2021*. Review of 108 articles from 2011–2020.

La Porta, R., Lopez-de-Silanes, F., Shleifer, A., & Vishny, R. (2000). Investor Protection and Corporate Governance. *Journal of Financial Economics*, 58(1–2), 3–27.

About the Author

Yunjung (Michelle) You, Ph.D. (Purdue University) is the founder of Apex Governance LLC and the designer of the Apex G-Score framework. A former faculty member at Minnesota State University, Mankato, she has over a decade of direct experience as both a minority and controlling shareholder in Korean listed and private companies—actively exercising shareholder rights, analyzing board structures, and observing how governance failures translate into value destruction. The Apex G-Score framework is grounded in this firsthand operational understanding of Korean corporate governance.

Legal Disclaimer

The author and Apex Governance LLC may hold equity positions in companies covered by the G-Score. However, all scores are computed through pre-defined, automated algorithms; individual holdings do not influence scoring outcomes.

The scoring methodology, including specific indicator definitions, item-level scoring rules, point allocations, archetype classification algorithms, and Kill Switch trigger conditions, constitutes proprietary intellectual property of Apex Governance LLC and is protected as trade secrets under applicable law.

Apex Governance LLC assumes no liability for any investment decisions made based on information contained in this document. All investment decisions and their consequences are the sole responsibility of the investor.