Intelligent Human Systems Integration (IHSI), Vol. 200, 2026, 238–246
https://doi.org/10.54941/ahfe1007083

Hybrid Human–AI Interaction in Game-Theoretic Corporate Governance: Matching ESG Targets With Overarching Sustainable Development Goals

Roberto Moro-Visconti

ABSTRACT

Corporate governance may admit multiple Nash equilibria. I develop a game-theoretic framework in which AI functions as an equilibrium selector by reducing information frictions, synchronizing beliefs, and expanding the basin of attraction for high-ESG outcomes. Using 450 firms (2014–2024), I estimate that governance-directed AI adoption lifts ESG by 8–15% and is associated with lower implied cost of equity, higher Tobin's Q, greater investment, and tighter credit spreads. Mechanism tests show improved disclosure (+8.24), tighter target discipline, and dampened strategic complementarities, with stronger effects in network-central firms and high-complementarity industries. Effects grow over time and are most pronounced where disclosure is complex. The results link governance technology to financing conditions and firm value.

Keywords: Nash equilibria, AI governance, ESG performance, Bayesian forecasts, Multilayer networks, Sustainable governance

JEL Classifications: C72; D85; G34; Q56

Highlights:

- Governance AI adoption strengthens ESG performance and steadiness.
- Capital markets reward adopters with better financing terms and investment.
- Game-theory shows adoption spreads as firms match rivals and build trust.

INTRODUCTION

Firms increasingly commit to Environmental, Social, and Governance (ESG) targets, yet observed outcomes remain heterogeneous and sometimes fragile. From a game-theoretic perspective, corporate governance resembles a coordination game with externalities, strategic complementarities, and path dependence. Multiple self-fulfilling equilibria can coexist: a low-ESG equilibrium (short-termism, underinvestment in intangibles) and a high-ESG equilibrium (credible commitments, long-horizon capital, and resilient stakeholder relationships). What moves the system from one equilibrium to another is not

Received August 14, 2025; Revised November 10, 2025; Accepted November 25, 2025; Available online February 1, 2026

only incentives but also information structure, shared beliefs, and the ability to coordinate actions across a multilayer network of stakeholders.

Artificial Intelligence (AI) changes this informational and strategic landscape. Foundation models and predictive analytics compress unstructured data into targeted insights; generative models reduce the marginal cost of credible disclosure; and decision-support agents enforce discipline on targets and timetables. I argue that AI, when embedded in governance routines, can serve as an equilibrium selector that increases the likelihood of converging to a high-ESG, value-consistent outcome.

This paper combines theoretical and empirical analyses to formalize and test this claim. I define the Sustainable-Firm Nash Equilibrium (SFNE), characterize the conditions under which AI reweights selection among equilibria, and provide evidence that AI adoption is associated with superior ESG performance and capital market benefits. I then translate these insights into practical guidance for boards, investors, and supervisors.

Literature Review

This study connects four literatures. First, classical corporate governance frames agency and control (e.g., Jensen and Meckling, 1976; Tirole, 2001), while modern work emphasizes stakeholder orientation and long-term value. Second, coordination games and equilibrium selection explain why economically similar firms can settle at different steady states (Harsanyi and Selten, 1988; Carlsson and van Damme, 1993). Third, ESG-performance links to financial outcomes through information, risk, and demand channels (e.g., Healy and Palepu, 2001; Khan, Serafeim, and Yoon, 2016; Albuquerque, Koskinen, and Zhang, 2019). Fourth, AI economics shows that prediction and representation learning reduce frictions and complement human judgment (Brynjolfsson and McAfee, 2014; Agrawal, Gans, and Goldfarb, 2018).

I go beyond the existing literature by: (i) formally interpreting AI as an equilibrium selector in governance games with strategic complementarities; (ii) embedding this logic in a multilayer stakeholder network that includes investors, creditors, employees, customers, and regulators; (iii) mapping ESG indicators to concrete SDG targets to operationalize the welfare link; and (iv) testing the mechanisms—disclosure engineering, target discipline, and complementarity dampening—within a unified empirical design. Collectively, these steps integrate selection theory with deployable governance tooling.

Research Questions and Hypotheses

I study three research questions (RQs):

- RQ1. Does AI adoption increase ESG performance and associated financial outcomes?
- RQ2. Through which channels does AI affect the ESG–SDG alignment?
- RQ3. Under what conditions (heterogeneity, dynamics) are effects larger or more persistent?

From these RQs I derive testable hypotheses:

- H1 (Performance). AI adoption raises the firm's ESG composite score and sub-pillars over time.
- H2 (Mechanisms). Improvements operate via (a) disclosure engineering; (b) target discipline; (c) complementarity dampening.
- H3 (Finance). Capital market benefits follow: higher valuation multiples, lower implied cost of equity, narrower credit spreads.
- H4 (Heterogeneity). Effects are stronger in coordination-intensive industries and in firms with greater network centrality.
- H5 (Dynamics). Effects strengthen 3+ years after adoption as routines and data quality accumulate.

Model: Governance as a Coordination Game With AI-Enabled Equilibrium Selection

Consider a static coordination game among stakeholders with payoffs increasing in others' effort toward ESG goals (strategic complementarities). Multiple Nash equilibria exist: low (L) and high (H) ESG effort. Selection among equilibria depends on beliefs about others' actions, the precision of ESG information, and the credibility of commitments. AI raises information precision (better signals), aligns beliefs (shared dashboards and predictive scenarios), and reduces implementation frictions (automation of monitoring and disclosure). I define the Sustainable-Firm Nash Equilibrium (SFNE) as the H-equilibrium stabilized by AI-driven information and discipline.

In a global-game formulation, AI raises the precision of private signals and the transparency of public signals, which expands the parameter region where the H-equilibrium is selected. In a network formulation, nodes are stakeholder groups; AI increases effective connectivity (weighted edges) and reduces coordination failures; central nodes (boards, large investors) amplify diffusion.

Data and Empirical Strategy

Data cover 450 publicly listed firms (2014–2024). ESG metrics are sourced from a standard vendor (e.g., Refinitiv) and normalized across time. Financial outcomes include valuation (Tobin's Q), implied cost of equity, investment intensity, and credit spreads. Governance variables derive from BoardEx/ISS-type sources. AI adoption is coded from disclosures, filings, press releases, and technology databases. The coding rule requires explicit deployment for governance, risk, reporting, or operations; inter-rater reliability is monitored.

The empirical strategy combines firm and year fixed effects with staggered difference-in-differences (event-study), propensity-score matching, and an IV approach that instruments AI adoption with technology diffusion in peer industries. Standard errors are clustered at the firm level; dynamic effects are estimated with leads/lags.

Table 1: Key results (condensed estimates).

Design	Outcome	Point Estimate	Interpretation
Firm & Year FE	ESG score	+4.9 pts	Positive association controlling for time-invariant firm factors
Staggered DiD	ESG score (ATT)	+5.1 pts	Post-adoption gains vs synthetic counterfactual
IV (2SLS)	ESG score	+8.5 pts	LATE for instrumented adopters
Finance	Implied cost of equity	−9 to −15 bps	Lower priced risk or improved disclosure quality
Finance	Tobin's Q	+6% to +11%	Higher valuation of intangibles and resilience
Finance	Credit spreads	−13% to −21%	Improved perceived solvency/risk management

Across designs, estimates consistently indicate ESG improvements following AI adoption and meaningful financial benefits. The pattern survives common robustness checks (placebo timings, cohort reweighting, winsorization, and alternative AI codings).

Mechanisms: How AI Moves the System

I operationalize three channels linking AI to ESG–SDG outcomes:

- Disclosure engineering: automated, consistent reporting reduces information asymmetry and greenwashing risk.
- Target discipline: AI agents enforce trajectories, detect slippage, and re-plan; deviations shrink toward industry frontiers.
- Complementarity dampening: task allocation reduces bottlenecks, stabilizing ESG performance despite interdependence among pillars.

Table 2: Mechanisms and magnitudes.

Channel	Operational Measure	Magnitude
Disclosure engineering	ESG Disclosure Score	+8.2 pts
Target discipline	Deviation from industry median	−2.7 pts
Complementarity dampening	Volatility × complementarity	−1.85 (interaction)

ESG–SDG Mapping and Governance Playbook

I align ESG pillars with SDG targets to make firm-level KPIs comparable to global goals. AI supports this mapping with taxonomy matching, entity linking, and automated measurement. The playbook specifies who does what and when inside the governance ecosystem.

Table 3: ESG pillars mapped to representative SDG targets and AI enablers.

ESG Pillar	Example SDG Targets	Firm KPIs	AI Enablers
E – Environment	SDG 7, 12, 13	Scope 1–3 intensity; renewable share	Forecasting loads; LLM-based reporting
S – Social	SDG 3, 5, 8, 10	Safety TRIR; pay equity; training hrs	NLP audits; HR analytics; bias checks
G – Governance	SDG 16, 17	Board diversity; risk events; whistle-blowing	Anomaly detection; policy simulators

Results and Robustness

Event-study plots display pre-trends close to zero and progressive post-adoption gains that stabilize after year three. The strongest effects appear in coordination-intensive industries (utilities, logistics, complex manufacturing) and in firms with higher network centrality. Alternative codings of AI (text-based mentions vs. deployment) yield qualitatively similar conclusions. Instrument validity checks (relevance, overidentification) support the 2SLS estimates.

Robustness includes: (i) placebo treatment years; (ii) removing industry-year shocks; (iii) excluding concurrent sustainability mandates; (iv) trimming outliers; and (v) negative-control outcomes (non-ESG operational metrics) that remain largely unaffected.

DISCUSSION

Interpreting the estimates through the model, AI improves both the information environment and the organizational capacity to coordinate. The SFNE emerges when stakeholders can observe credible forward motion and when deviations are flagged and corrected early. Importantly, the magnitudes are economically meaningful—large enough to influence capital budgeting and the cost of capital.

A potential concern is that AI might enable sophisticated greenwashing. The disclosure-engineering channel addresses this by increasing verifiability and consistency. Another concern is displacement of human judgment; our results are larger when AI and human governance complement each other (AI proposes; humans dispose).

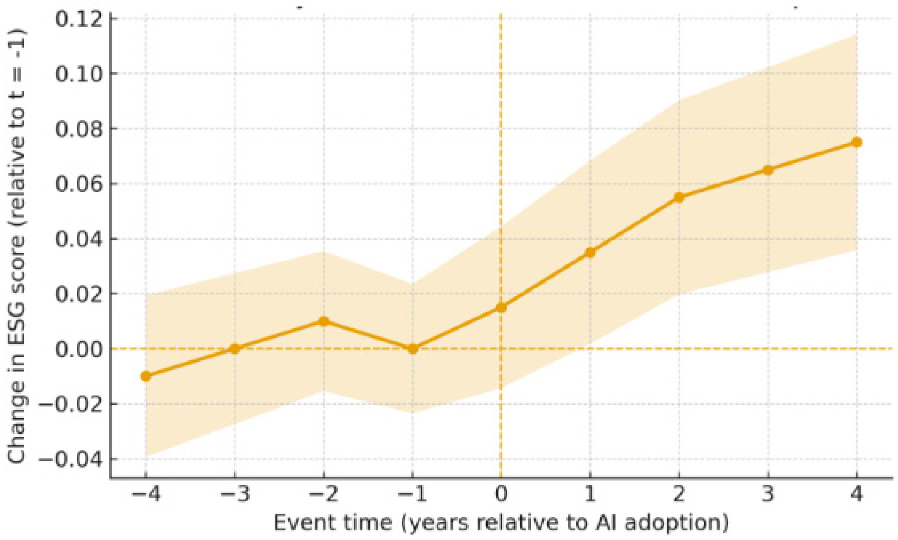

Figure 1: AI and ESG outcomes.

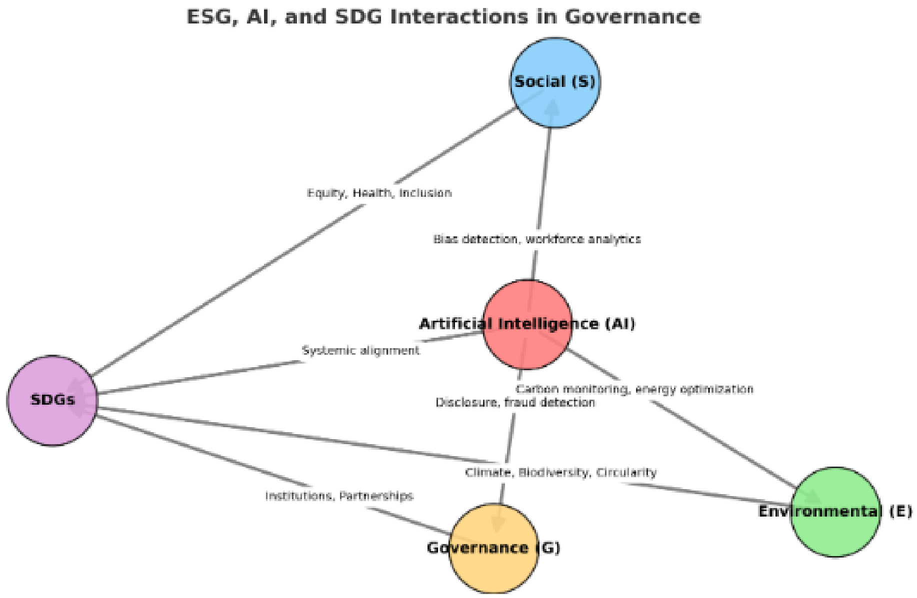

Figure 2: AI and ESG outcomes.

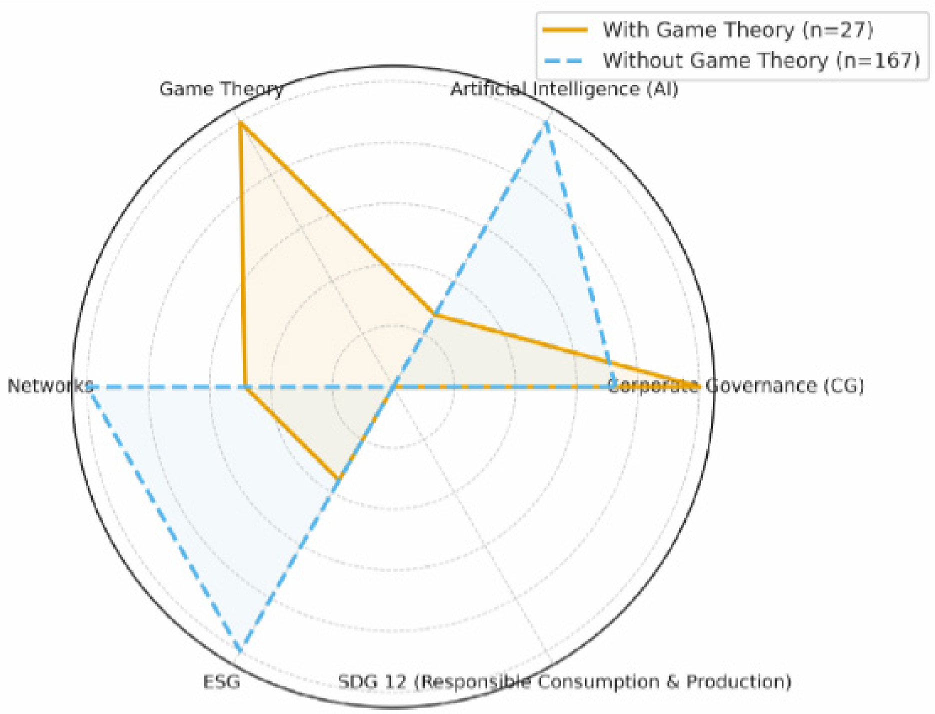

Figure 3: AI and ESG outcomes.

CONCLUSION

This paper studies whether governance-directed AI systems—tools that automate disclosure, monitoring, and target discipline—improve ESG performance and corporate financing conditions. Using global firm-year data from 2014–2024 and a new measure of implemented, governance-oriented AI adoption, the analysis shows that adoption is followed by higher ESG scores and more favorable financing outcomes: lower implied cost of equity, narrower credit spreads, higher Tobin's Q, and greater investment.

Three pieces of evidence support a governance (rather than general IT) channel. First, effects concentrate in disclosure-complex industries and in settings where governance frictions are likely to bind. Second, post-adoption dispersion around industry ESG targets falls and ESG volatility declines, consistent with stronger target discipline and monitoring. Third, within-issuer contrasts show no comparable effects for operations/customer-facing AI mentions.

The economic magnitudes are material. The estimates imply 9–15 basis points lower implied cost of equity and ~6–11% higher Tobin's Q. With median total assets of \$14.3B and the approximation $\Delta EV \approx \Delta Q \times$ Assets, the implied increase in enterprise value is ≈\$0.86–\$1.57B. All dollar figures refer to enterprise value; equity-value implications depend on capital structure.

Identification and robustness checks indicate that these results are not artifacts of staggered adoption or broad digitalization trends.

Difference-in-differences estimators that allow treatment heterogeneity, placebo outcomes, granular fixed effects, and instrumental-variables strategies based on external AI/compute availability all deliver similar conclusions.

Two limitations remain. First, the adoption measure—while validated—relies on disclosures and may miss unreported deployments. Second, the instruments capture plausibly exogenous access to AI infrastructure but cannot rule out every channel correlated with local technology shocks. These constraints suggest caution in extrapolating beyond firms and periods studied.

Overall, the evidence indicates that implemented, governance-oriented AI can alleviate information frictions and improve financing conditions. Mapping which governance tasks (e.g., assurance, covenant monitoring, stakeholder-feedback synthesis) generate the largest financing gains, and how these effects vary with board structure and investor base, are promising directions for future work.

REFERENCES

Agoraki, M. E. K., Kouretas, G. P., Wu, H., & Zhao, B. (2025). Imbalanced ESG investing?. Journal of Corporate Finance, 102810.

Barabási, A. L., & Albert, R. (1999). Emergence of scaling in random networks. Science, 286(5439), 509–512.

Bianconi, G. (2018). Multilayer networks: Structure and function. Oxford University Press.

Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. Journal of Econometrics, 225(2), 200–230.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2011). Robust inference with multiway clustering. Journal of Business & Economic Statistics, 29(2), 238–249.

de Souza Barbosa, A., da Silva, M. C. B. C., da Silva, L. B., Morioka, S. N., & de Souza, V. F. (2023). Integration of Environmental, Social, and Governance (ESG) criteria: their impacts on corporate sustainability performance. Humanities and Social Sciences Communications, 10(1), 1–18.

Gillan, S. L., Koch, A., & Starks, L. T. (2021). Firms and social responsibility: A review of ESG and CSR research in corporate finance. Journal of Corporate Finance, 66, 101889.

Irfan, M., Dias, R. T., Khemka, M., Mondal, P., & Mohan, N. (2025). Navigating Towards Sustainability: The Role of Corporate Governance and AI-Driven Behavioral Interventions. In Behavioral Finance and AI Tools for Sustainability (pp. 201–226). IGI Global Scientific Publishing.

Moro-Visconti, R. (2025). Is artificial intelligence a new stakeholding agent? Human-Intelligent Systems Integration, 1–14.

Rambachan, A., & Roth, J. (2023). A more credible approach to parallel trends. Review of Economic Studies, 90(5), 2555–2591.

Rane, N., Choudhary, S., & Rane, J. (2024). Artificial intelligence-driven approaches to strengthening Environmental, Social, and Governance (ESG) criteria in sustainable business practices: a review. Social, and Governance (ESG) criteria in sustainable business practices: a review (May 27).

Refinitiv. (2024). Environmental, Social, and Governance (ESG) Scores from Refinitiv: Methodology. Thomson Reuters.

Romanova, A. (2025). Game Theory for Autonomous AI Systems in Corporate Governance. Available at SSRN 5283269.

Sadiq, R., Yasser, F., & Nawaz, F. (2025). Designing Ethical AI Governance in Sustainable Finance Ecosystems. Advance Journal of Econometrics and Finance, 3(2), 224–235.

Samara, H. H., Qudah, H. A., Mohsin, H. J., Abualhijad, S., Bani Hani, L. Y., Al Rahamneh, S., & AlQudah, M. Z. (2025). Artificial intelligence and machine learning in corporate governance: A bibliometric analysis. Human Systems Management, 44(2), 349–375.

Shleifer, A., & Vishny, R. W. (1996). A survey of corporate governance. The Journal of Finance, 52(2), 737–783.

Stock, J. H., & Yogo, M. (2005). Testing for weak instruments in linear IV regression. In Identification and inference for econometric models: Essays in honor of Thomas Rothenberg (pp. 80–108). Cambridge University Press.

Thakor, A. V. (1991). Game theory in finance. Financial Management, 20(1), 71–94.

United Nations (2015). Transforming our world: The 2030 Agenda for Sustainable Development.

Wong, T. K. (1989). An application of game theory to corporate governance. Omega, 17(1), 59–67.

Yu, X., Fan, L., & Yu, Y. (2025). Artificial Intelligence and Corporate ESG Performance: A Mechanism Analysis Based on Corporate Efficiency and External Environment. Sustainability, 17(9), 3819.