ISSN 0956-8549

Artificial Intelligence in the Boardroom

By

Daniel Ferreira
Jin Li

FINANCIAL MARKETS GROUP DISCUSSION PAPER NO. 958

June 2026

Any opinions expressed here are those of the authors and not necessarily those of the FMG. The research findings reported in this paper are the result of the independent research of the authors and do not necessarily reflect the views of the LSE.

Artificial Intelligence in the Boardroom*

Daniel Ferreira Jin Li

April 5, 2026

Abstract

We analyze how the use of artificial intelligence affects the monitoring and advisory relationship between CEOs and corporate boards. AI can serve as a private advisor to CEOs, partially substituting for board advice, and thus reducing CEOs' incentives to share information with directors. In equilibrium, firms respond by reducing board independence to restore information flows. Lower monitoring intensity decreases CEO turnover, making CEOs more entrenched. Lower dismissal risk allows firms to reduce CEO compensation. We show that AI adoption by the board mitigates some of the negative aspects of the CEO's AI use, but cannot restore efficient levels of board monitoring. Our analysis suggests that the adoption of AI in the boardroom may have unintended consequences for corporate governance.

Keywords: Boards, CEOs, Artificial Intelligence, Corporate Governance

---

[*] We thank Mike Burkart, Felix Zhiyu Feng, Luis Garicano, Dirk Jenter, Liyan Yang, and seminar participants at EDHEC Business School, LSE, and the U. of Delaware Corporate Governance Symposium. We thank Zehao Zhang for excellent research assistance. Ferreira: London School of Economics, CEPR, and ECGI. Li: HKU Business School.

1 Introduction

Artificial intelligence systems have evolved from sophisticated chatbots to agents capable of solving problems and performing complex tasks in a semi-independent manner. Such systems can both help experts and replace them.[1] This transformation has been particularly pronounced in domains requiring analytical reasoning, pattern recognition, and strategic assessment—capabilities that were once the exclusive province of human expertise.[2] Major corporations are already deploying AI systems that can analyze vast datasets, identify market trends, and evaluate strategic options. The speed and scope of this technological shift raise questions about how organizations structure their decision-making processes and divide tasks between humans and artificial intelligence systems.[3]

In this paper, we develop a framework for studying the impact of generative artificial intelligence on the relationship between CEOs and corporate boards. As AI increasingly serves as an analytical tool for executive decision-making,[4] it potentially disrupts the traditional information flows between executives and directors. While CEOs and other executives can now access AI-generated insights for strategic planning, idea generation, and project assessment, this capability may reduce their reliance on board expertise and alter their incentives to share information.[5] We show that these changes in communication dynamics have important implications for board composition, monitoring, and executive

---

[1] Korinek (2025) shows how AI systems can enhance the productivity of economics researchers by both working in collaboration with them and carrying out independent research tasks.

[2] In a randomized controlled trial with 758 Boston Consulting Group consultants, Dell'Acqua et al. (2023) find that those using GPT-4 completed 12.2% more tasks, finished them 25.1% faster, and achieved over 40% higher quality output compared to controls.

[3] Babina et al. (2025) use resume and job posting data to document significant flattening of hierarchical structures with AI adoption, showing increases in junior-level workers and decreases in middle management roles.

[4] In a large international survey, Yotzov et al. (2026) find that AI adoption by top executives increased by more than 50% in 2025, and that CEOs are more frequent users of AI than other executives. Kenny, Kowalkiewicz, and Oosthuizen (2025) present cases and examples of how CEOs can use AI for strategic planning. Csaszar, Ketkar, and Kim (2024) present evidence of the use of AI as a strategic decision-making tool in firms.

[5] Larcker et al. (2025) argue that AI has the potential to alter the relationship between executives and directors significantly. Shekshnia and Yakubovich (2025) discuss how AI helps directors.

compensation.

The model is as follows. A CEO exerts effort to identify a problem (e.g., find a deal, generate an idea, etc.). Identifying a problem is a "messy job" (Garicano, Li, and Wu (2026)), that is, a job that must be performed by humans. Once a problem is identified, the CEO must assess it and decide whether to address it on her own or with the board's help. The CEO has access to an AI agent. We consider AI primarily as an oracle with a dual role: it is both a tool and a coworker.[6] Ransbotham et al. (2025) argue that "*this tool-coworker duality breaks down traditional management logic, which assumes that technology either substitutes or complements, automates or augments, is labor or capital, or is a tool or a worker, but not all at once*."[7] Furthermore, what makes AI distinctive in this regard is that its consultation is private and nearly costless, setting it apart from traditional sources of outside advice. To model this duality, we assume that the AI technology has two capabilities: it enhances the CEO's problem-solving skills (a *skill-augmenting capability*), and it also solves some problems independently of the CEO's ability (a *problem-solving capability*).[8]

We consider the problem of a group of controlling shareholders who must choose the board's monitoring intensity (i.e., board independence) and the CEO's compensation contract, which we jointly refer to as a *governance structure*. An optimal governance structure must incentivize the CEO to exert effort to identify deals and to ask the board for help when the CEO is unable to solve a problem on her own. Asking for help increases the likelihood that the board replaces the CEO because it is perceived as a negative signal about the CEO's ability. To induce the CEO to communicate truthfully with the board, the optimal governance structure must involve a board that is partially "CEO-friendly," that is, one that may retain the CEO even when a better replacement is available. As a conse-

---

[6] An oracle "*serves as an advisor, offering information, predictions, or recommendations while leaving all decisions and actions to human users*" (Jiang (2025), p. 114).

[7] In a survey on the use of agentic AI in business decision-making, Ransbotham et al. (2025) find that 76% of executives view agentic AI more as a coworker than a tool.

[8] This dual role of AI features prominently in recent economic analyses of the impact of AI on the future of work. For example, Acemoglu, Autor, and Johnson (2026) call these roles "collaboration" and "automation;" Demirer et al. (2026) distinguish between "worker augmentation" and "task automation."

quence, the optimal structure will typically commit to some degree of ex post inefficient CEO entrenchment to enhance the CEO's ex ante incentives to seek the board's advice.

A CEO with access to an AI agent with problem-solving capabilities can use it as a substitute for the board's advice. The more capable AI is, the less likely is the CEO to ask the board for help. Therefore, to restore information flows between the CEO and the board, the optimal board structure must be more CEO-friendly when such a capability exists (i.e., when AI is a coworker). As the CEO faces a lower risk of dismissal under a friendlier board, the optimal compensation level is lower with AI than without it. Because the CEO's incentive constraints must be binding, the decrease in dismissal risk is exactly offset by the reduction in pay, leaving the CEO no better off. Thus, the AI-driven increase in CEO entrenchment destroys firm value, without benefiting the CEO.

In contrast, if AI augments CEOs' skills (i.e., CEOs use AI primarily as a tool), CEOs become more productive, and firms find it easier to incentivize them to exert effort. Thus, CEOs can be induced to exert effort with lower pay. However, lower pay reduces CEO incentives to communicate with the board. Thus, as in the case in which AI serves as a coworker, the optimal governance structure requires low pay and low monitoring intensity.

We conclude that CEO access to AI increases board friendliness and reduces CEO pay, regardless of whether CEOs use AI primarily as a tool or as a coworker. The value implications of AI adoption by CEOs are ambiguous: while AI's problem-solving capabilities are welfare-destroying, its skill-augmenting capabilities are welfare-increasing.

While CEO AI adoption unambiguously reduces both the board's monitoring intensity and CEO pay levels, a natural follow-up question is whether board AI adoption can counter the effects of CEO AI adoption. We extend the model to allow the board to directly benefit from AI adoption in its advisory and monitoring roles. We show that board AI adoption attenuates some of the effects of CEO AI adoption on monitoring and pay levels. However, to reverse the original effects, shocks to AI technology must be significantly biased toward the board, meaning AI must augment directors' skills more than

CEOs'. If we instead expect AI to benefit CEOs and boards in similar ways (or to benefit the CEO more), AI adoption must decrease board monitoring and CEO pay levels.

Our paper is the first formal analysis of the impact of AI on corporate boards. Our model relates to a theoretical literature on the dual role of boards as advisors and monitors of management,[9] which includes Adams and Ferreira (2007), Harris and Raviv (2008), Baldenius, Melumad, and Meng (2014), Baldenius, Meng, and Qiu (2019), Levit (2020), Drymiotes and Sivaramakrishnan (2021), and Jiang and Laux (2023), among others.[10] More generally, the paper relates to the broader literature on information asymmetries in the boardroom, as in Warther (1998), Raheja (2005), Song and Thakor (2006), Baranchuk and Dybvig (2009), Inderst and Mueller (2010), Levit (2012), Malenko (2014), Levit and Malenko (2016), Chakraborty and Yilmaz (2017), Chemmanur and Fedaseyeu (2018), and Donaldson, Malenko, and Piacentino (2020).[11]

Our model follows Adams and Ferreira (2007) in that the CEO must first provide information to the board before receiving advice. That is, ours is a "two-way" communication model. In this setup, CEO-friendly boards can be optimal: the board commits to low-intensity monitoring to provide the CEO with incentives to share information. The logic is reminiscent of Burkart, Gromb, and Panunzi (1997) and Aghion and Tirole (1997), who show that ex-post intervention by principals undermines agents' ex-ante incentives.[12]

As in Song and Thakor (2006), in our model, the CEO is responsible for generating a

---

[9] A more recent literature has also focused on the "mediating" role of boards. Burkart, Miglietta, and Ostergaard (2023) present evidence that boards monitor management and mediate between different types of shareholders. Ewens and Malenko (2024) provide evidence that VC-backed startup boards advise and mediate between investors and executives. Also consistent with a mediating role of boards, Ferreira, Ferreira, and Mariano (2018) show evidence that changes in creditors' power relative to shareholders affect appointments to the board.

[10] To the best of our knowledge, the advisory role of boards is first mentioned in the economics and finance literature by Fama and Jensen (1983), who argue that directors "*provide an important support function to the top managers in dealing with specialized decision problems*" (p. 314).

[11] See also Malenko (2024) for a comprehensive review of the literature on information flows in corporate governance.

[12] More generally, our model relates to the theoretical literature on delegation (e.g., Holmström (1984), Melumad and Shibano (1991), Dessein (2002), Alonso and Matouschek (2008), and Amador and Bagwell (2013), among others) .

project idea with an unknown probability of success before deciding whether to reveal it to the board. CEOs differ in ability, enjoy private benefits, and face career concerns. The board monitors by gathering information about the CEO and occasionally replacing a low-ability CEO. CEO compensation is endogenous and depends on the CEO's performance and the board's belief about the CEO's type. Thus, unlike Adams and Ferreira (2007), our model has implications for CEO turnover and compensation.

A key innovation of our paper is to model the relationship between a CEO and a board as a knowledge hierarchy, as in Garicano (2000) and Garicano and Rossi-Hansberg (2006). Following the original contribution of Ide and Talamàs (2025), we introduce AI within a knowledge-hierarchy setup. As in Nikolowa (2015), agents learn about their knowledge over time. As in Fuchs, Garicano, and Rayo (2015), there is asymmetric information about agents' knowledge.

Our analysis shows that AI adoption by CEOs and boards can have unintended consequences for corporate governance due to the interactions between the board's monitoring and advising roles. For evidence that boards perform an important advisory role (which at times conflicts with its monitoring role), see Coles, Daniel, and Naveen (2008), Linck, Netter, and Yang (2008), Faleye, Hoitash, and Hoitash (2011), Field, Lowry, and Mkrtchyan (2013), Harford and Schonlau (2013), Adams, Akyol, and Verwijmeren (2018), de Haas, Ferreira, and Kirchmaier (2021), and Ewens and Malenko (2024).

There are a few theoretical papers on AI in organizations and governance. An early contribution is Athey et al. (2020), who introduces an AI agent to Aghion and Tirole's model of authority. Zhong (2025) develops a framework to study the interactions between humans and technologies in a multi-layer decision-making problem. His model implies that the optimal human-AI authority allocation depends on the trade-off between correcting errors and introducing new ones, with AI typically best as the final decision-maker. Cheng, Dogan, and Yildirim (2025) study the optimal replacement of workers by AI agents in a team production problem. Huang and Ouyang (2025) develop a cheap-talk model of AI-based advising in which AI agents are unbiased (unlike human agents) but face difficulties in codifying soft information. More closely related to our work is Chen and Han (2026). They develop a model in which biased CEOs may have access to AI technology. Because AI increases the granularity and precision of hard information available to CEOs, there is more scope for biased decisions, making agency problems more severe. Firms can mitigate some of these negative effects by changing contracts, but these fixes come at a cost. They conclude that AI adoption may reduce profits and welfare.

Finally, our paper connects to recent empirical work documenting AI's unintended consequences on worker behavior. Cowgill, Hernandez-Lagos, and Wright (2024) show that generative AI reduces screening accuracy by 4-9% as less-qualified individuals use it to mimic expert communication, thereby fundamentally altering how information flows between applicants and evaluators. Chen et al. (2025) show that AI adoption leads to lower worker motivation. These empirical findings underscore the importance of AI as a source of information cost. Our model shows that firms can partially mitigate these adverse effects of AI adoption by adjusting their governance structures.

We present our model setup in Section 2. Section 3 shows the impact of AI on board learning. Section 4 solves the model in a pre-AI economy. Our main results are then presented in Section 5. Section 6 extends the model to allow for AI adoption by the board. Section 7 concludes. All omitted proofs are in the Appendix.

2 Model Setup

We present a model of the relationship between a CEO and a board of directors. Following Ide and Talamàs (2025), the model builds on Garicano's (2000) knowledge hierarchy framework, augmented by AI. In this framework, the CEO plays the role of a "worker" who identifies and assesses investment opportunities, while the board serves as an "advisor" who can help the CEO when needed. In addition, the board also monitors the CEO. We begin by describing the agents, the technology, and the information structure. We then present the full game in detail.

Agents, Technology, and AI. We consider a firm owned by one or multiple shareholders, who are residual claimants of the firm's cash flow. There are two periods. The firm must hire one CEO and one director in each period. We set the board size to one. Shareholders have no specialized human capital and, thus, cannot become CEOs or directors. All agents—CEOs, directors, and shareholders—are risk-neutral, have zero discount rates, have zero outside utility, and are protected by limited liability.

The CEO's job is a "messy job:" it is an indivisible bundle of tasks that cannot be automated, involving both hard and soft skills (Garicano, Li, and Wu (2026)). Specifically, the CEO operates a technology represented by a continuum of *problems* (i.e., potential deals, projects, ideas, etc.). If the CEO exerts effort at a fixed personal cost of $c > 0$, she identifies a problem of unknown *difficulty* $x \sim U[0, 1]$. One interpretation is that c is the CEO's personal cost of "deal sourcing." We initially assume that only CEOs can identify deals; this task cannot be facilitated or performed by AI agents. We relax this assumption in Subsection 5.3. From now on, we refer to "deals" and "problems" interchangeably.

After identifying a deal, the CEO must determine whether it creates value for shareholders; x is the (deal-specific) difficulty in making such an assessment. To tackle this problem, the CEO relies on her *knowledge*, which is indexed by $z \sim U[0, 1]$. Knowledge is persistent. If $x \leq z$, a CEO who spends one unit of time on the problem can solve it. We set the value of each solved problem to 1. If $x > z$, the CEO cannot solve the problem, in which case the output is zero.

If, at the beginning of a period, the CEO identifies a problem she cannot solve, she can ask the board for help. That is, boards serve as advisors who specialize in providing help (as in Garicano (2000), Garicano and Rossi-Hansberg (2006), and Ide and Talamàs (2025), among others). Thus, firms are structured as two-layer knowledge hierarchies. Board directors have knowledge equal to one; that is, they can solve problems of any difficulty level. This assumption is stronger than necessary; we only need the board's expected knowledge to be sufficiently high.

The board's time and attention are constrained. If the CEO asks the board for help, the board allocates less time and attention to routine "compliance tasks." Specifically, diverting the board's attention away from the compliance tasks destroys one unit of output with probability $\eta \in (0, 1)$. That is, if the board spends one unit of time helping the CEO assess a deal, the expected output is $1 - \eta$. Thus, if the CEO can perform the assessment herself, it is best not to involve the board. Henceforth, we refer to $1 - \eta$ as the *board's problem-solving capability*.[13]

The CEO also has access to an AI agent. The firm cannot deny the CEO access to AI; CEOs' AI adoption decisions are private information. We represent the capabilities of the AI agent by a pair of parameters $(\alpha, \kappa) \in [1, \overline{\alpha}] \times [0, 1]$. The AI agent has two capabilities. First, it serves as a tool to help the CEO solve problems. If the CEO uses AI as a tool, she can solve any problem that has difficulty $x \leq \alpha z$. That is, parameter α measures how much AI enhances the CEO's ability to solve problems. We call α AI's *skill-augmenting capability*. Conceptually, AI's skill-augmenting capability is no different from other technological innovations—such as computers—that help humans perform complex tasks and solve more problems.[14]

Second, AI agents can solve some problems autonomously with minimal human input. That is, AI is not only a tool, but also a coworker. An AI model with *problem-solving capability* κ can autonomously solve problems with difficulty $x \leq \sqrt{\kappa}$. That is, $\sqrt{\kappa}$ is the AI agent's knowledge, which implies that κ is the probability that AI can solve a problem conditional on the CEO's not being able to solve it: $\kappa = \Pr[x \leq \sqrt{\kappa} \mid x > \alpha z]$. AI's autonomous problem-solving capability is what arguably differentiates AI models from other innovations. While α represents AI's ability to complement an agent's knowledge, κ represents AI's ability to be a substitute for human decision-making. Thus, the pair of parameters (α, κ) captures the idea that the CEO can view AI as both a tool and a coworker

---

[13] Parameter η plays the role of the "cost of help" in standard knowledge hierarchy models (e.g., Ide and Talamàs (2025)).

[14] Yotzov et al. (2026) show evidence that the most common types of AI technologies used by businesses are data processing with machine learning and text generation with LLMs. These are both examples of skill-augmenting capabilities.

(Ransbotham et al. (2025)).[15]

For most of the paper, parameters (α, κ) fully describe AI's capabilities. After presenting our core model, we also consider three natural extensions: AI may lower the cost of deal sourcing (Subsection 5.3), AI may improve the board's advice (Subsection 6.1), and AI may enhance the board's monitoring ability (Subsection 6.2).

Information. Under complete information about knowledge levels and problem difficulty, a CEO with access to an AI agent with capabilities (α, κ) would proceed as follows. If she draws a problem with difficulty $x \leq \tilde{z} := \max\{\alpha z, \sqrt{\kappa}\}$, she can solve it without involving the board. If $x > \tilde{z}$, the CEO asks the board for help. The net effect of AI is identical to that of a technology that makes the CEO's "effective knowledge" equal to $\tilde{z}$. In other words, whether AI is used as a tool or a coworker does not matter.

Things are different when information is incomplete. From now on, we make the following assumptions about the information environment.

Assumption 1 (Information Environment). *Let z denote the CEO's knowledge and x denote the difficulty of a problem identified by the CEO. We assume that:*

(i) No one observes z or x;

(ii) CEOs learn whether they can solve a problem immediately after drawing it;

(iii) No one knows with certainty whether an AI agent can solve a specific problem.

Assumption 1(i) implies that all agents are symmetrically uninformed at the beginning of the first period. Assumption 1(ii) (which follows Nikolowa (2015)) implies that CEOs

---

[15] The active use of AI agents to analyze complex business decisions independently is still rare at the time of this paper's writing, but AI agents are already being deployed in business to independently tackle more specialized tasks such as insurance claims processing, inventory management, and financial trading. Survey evidence on the use of agentic AI in business decision-making shows that even among the most advanced adopters, only 54% of respondents are pursuing scenarios in which AI decides and implements autonomously (Ransbotham et al. (2025)); yet the same respondents are two to three times more likely to expect AI systems to work independently from humans and have decision-making authority in three years compared to today.

must update their beliefs about their own knowledge after drawing a problem. Assumptions 1(i) and 1(ii) jointly imply that even if AI's problem-solving capability κ is known, the CEO still doesn't know whether AI can independently solve a problem, because the problem's difficulty is unknown.

Assumption 1(iii) implies that an AI agent always offers a solution when asked, even when it cannot solve the problem. In other words, AI models make mistakes, and it is not possible to know *with certainty* whether the proposed solution is correct before trying it out. This assumption is compatible with an AI agent knowing it makes mistakes and also knowing the probability that it cannot solve a problem. For example, we can reinterpret κ as an AI-provided "confidence score" in its answer. Alternatively, we can reinterpret κ as the CEO's confidence in the AI agent's answer after analyzing it. Assumption 1(iii) rules out only the possibility that the CEO knows with absolute certainty whether an AI-provided answer is objectively correct or incorrect.[16]

Our assumptions imply that when a CEO does not now how to assess a deal, she must decide whether to ask *either* AI *or* the board for help. Because the CEO cannot know whether the AI agent's solution is correct, the sequential strategy of first asking the AI agent and then the board leads to the same expected outcome as asking only the board if $1 - \eta > \kappa$ and only the AI agent if $1 - \eta \leq \kappa$.

Timing. Within a period, the timing of actions is as follows.

(i) Shareholders sign contracts with a CEO with (unobservable) knowledge $z \sim U[0, 1]$ and a board of (observable) *ex ante* type $\pi \in [0, 1]$. The CEO enjoys a private benefit $B > 0$.

(ii) The CEO decides whether to exert effort at a private cost c. If the CEO exerts effort, she finds a "deal" (i.e., a problem) with (unknown) difficulty $x \sim U[0, 1]$. The CEO

---

[16] It is certainly the case that, with the right prompt, an AI agent may admit it does not know the answer to a particular question, especially for a fully objective one, such as a math problem. However, on tasks for which there is no clear metric of correctness, such as management and strategy tasks, the human agent is likely to remain uncertain about the quality of an AI-generated solution (Garicano, Li, and Wu (2026)).

then learns her "*ex post* type," that is, whether $x \leq \alpha z$ (type g) or $x > \alpha z$ (type b). After learning her type, she chooses one of three actions: $m \in \{m_a, m_b, m_c\}$, where m_a denotes asking the AI agent to solve the problem autonomously, m_b denotes disclosing the deal to the board, and m_c denotes solving the problem by herself.

(iii) The state of the world, $\omega \in \{\omega^{tough}, \omega^{weak}\}$, is realized, where π (the board's *ex ante* type) denotes the probability of $\omega = \omega^{tough}$. The board then decides whether to fire or retain its incumbent CEO. If the board decides to retain the CEO, it commits to a bonus schedule for the next period.

(iv) Output $y_t \in \{0, 1\}$ is realized. The CEO earns a bonus $w_t \geq 0$ if $y_t = 1$.

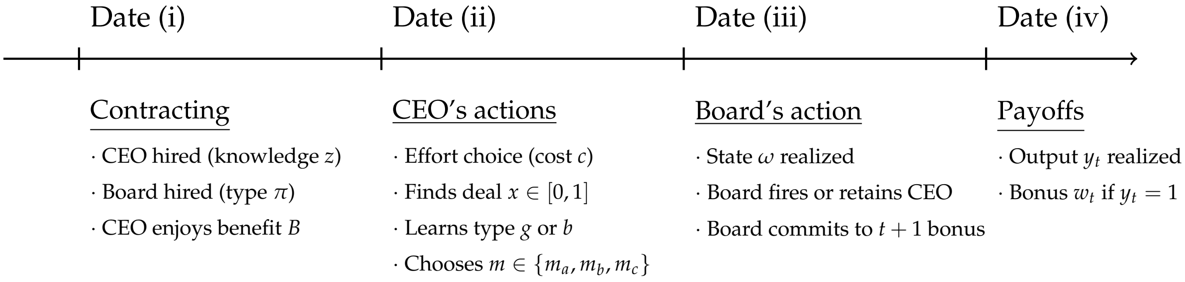

Figure 1: Timing of the game within a period.

Figure 1 provides a summary of the timing. We now provide a more detailed description of the game.

Board independence and monitoring. At Date (i), shareholders choose a board with *ex ante* type $\pi \in [0, 1]$, denoting its "independence" (i.e., its degree of alignment with shareholders' interests). Formally, we think of π as the result of an action that stochastically reduces the board's private cost of firing a CEO: with probability π, the board is "tough" (state of the world ω^{tough}) and faces no cost of firing the CEO, while with probability $1 - \pi$, the board is "weak" (state of the world ω^{weak}) and the cost of firing the CEO is prohibitively high. The state of the world, ω, is realized at Date (iii). Note that shareholders can commit not to fire a CEO by choosing $\pi = 0$. In contrast, the board cannot commit to always firing a CEO; even if $\pi = 1$, firing a CEO must still be incentive-compatible for the board. Thus, board monitoring is a real option to replace the CEO in state ω^{tough}.

Our interpretation is that π is determined by the board's structure and composition. A less independent board faces a greater cost of monitoring the CEO, and is thus less likely to replace an underperforming CEO.[17] As in Adams and Ferreira (2007), we call π the board's *monitoring intensity*. The interpretation is that CEOs regularly take actions to entrench themselves, such as co-opting board members. This entrenchment is a byproduct of the collaboration between the CEO and the board. By working closely with the CEO, the board becomes CEO-friendly. A friendlier board has lower monitoring intensity.

Contracts. In Date (i) of period $t = 1$, the firm offers a take-it-or-leave-it contract to a CEO with unknown knowledge $z \sim U[0,1]$. In Date (i) of period $t = 2$, the firm offers a take-it-or-leave-it contract to either a new CEO (again with knowledge $z \sim U[0,1]$) or the incumbent CEO (i.e., the same CEO as in $t = 1$), in which case the contract must be the one agreed on Date (iii) of period $t = 1$.[18] CEOs' outside options remain fixed at zero.[19] At the beginning of each period, an employed CEO enjoys a nonpecuniary private benefit of $B > 0$. CEOs have no external wealth. Firms can offer CEOs bonuses contingent on output $y_t \in \{0,1\}$ at the end of each period. Without loss of generality, we assume CEOs receive positive bonuses only when the output is high (i.e., $y_t = 1$). Let w_t denote the bonus at the end of period t conditional on $y_t = 1$, $t \in \{1,2\}$. We assume that the

---

[17] For simplicity, we assume that shareholders face no cost of choosing π; we could easily add a direct cost of independence without changing the nature of the problem.

[18] If the firm does not commit to a contract on Date (iii), the firm may use y_1 to update its beliefs about the incumbent CEO's type when setting her contract at $t = 2$. Our main results remain unchanged under this alternative timing, but the formal proofs are longer due to the need of specifying off-equilibrium-path beliefs that depend on y_1 realizations.

[19] Under a different information structure where competing firms could observe the CEO's decision to ask the board for help, in a separating equilibrium, CEOs who didn't ask for help would have better outside options. This would make asking for help a more negative signal, and inducing a CEO to ask the board for help would require an even lower monitoring intensity. This variation of the model is more complex but yields no significant further insights.

firm can commit to one-period contracts only. Thus, at $t = 1$, the firm cannot commit to bonuses that depend on the CEO's employment status at $t = 2$; this rules out contracts compensating CEOs for without-cause dismissal.[20] That is, at each new period, the firm and the CEO must negotiate a new contract. We also assume that the CEO can "credibly commit" to consuming any compensation received at the end of a period, for example, by borrowing against that amount. This assumption implies that the firm cannot recoup w_1 as part of the negotiation of the $t = 2$ contract, thus avoiding the introduction of lifetime contracts through the back door.[21]

In reality, CEO contracts are either fixed-term or "at will" (i.e., with no explicit termination date; see Gonzalez-Uribe and Groen-Xu (2017)). Fixed-term contracts usually last no longer than five years, but renewal is common. Most companies opt for at-will contracts, which do not provide protection against dismissal and are, in effect, a series of short-term contracts that are continuously renegotiated. We can interpret our one-period contracts either as fixed-term contracts with the possibility of renewal or as at-will contracts.[22]

CEO's effort and actions. At the beginning of Date (ii), if the CEO exerts effort at a cost of $c > 0$, she identifies a deal of difficulty x. The CEO does not observe x. Once the CEO identifies a deal, she still needs to spend time assessing it before deciding whether to invest. Immediately after identifying a deal, the CEO learns whether she can assess it herself (with the help of AI as a skill-augmenting technology). That is, the CEO learns

---

[20] We can allow for capped separation compensation without changing the results qualitatively. While compensation for without-cause dismissal is common in CEO contracts, it is subject to both hard and soft constraints. The US tax code heavily penalizes both executives and corporations for excessive severance payments, effectively limiting these payments to no more than three times the base salary. In addition, proxy advisors often have explicit guidelines flagging excessively generous separation contracts.

[21] The theoretical literature on delegation typically differs from optimal contracting and mechanism design approaches by restricting the space of feasible contingent transfers, often ruling out such contracts entirely. A notable exception is the recent work by Chen (2024), who provides a dynamic model of authority delegation under optimal compensation contracts. Similarly, our model combines features from both approaches.

[22] Our assumption of short or fixed-term contracts is realistic, but it is also with a loss of generality. At the cost of increased complexity, we could introduce alternative frictions that deliver similar qualitative results. However, the model would lose much of its simplicity and tractability.

that she is one of two *ex post* types: $x \leq \alpha z$ (type g) or $x > \alpha z$ (type b). After learning her type, she chooses one of three actions: $m \in \{m_a, m_b, m_c\}$. The board can observe $\hat{m} \in \{m_b, \overline{m}_b\}$ where $\overline{m}_b := \{m_a, m_c\}$. To allow for mixed strategies, let $P_t(\tau)$ denote the period-t probability distribution over m (i.e., the CEO's behavior strategy) conditional on the CEO's knowledge of her *ex post* type, $\tau \in \{b, g\}$. $P_t(\tau)$ determines $p_t(\tau) \in [0, 1]$, which is the probability that a CEO of *ex post* type τ in period t asks the board for advice, i.e., the probability of $m = m_b$. If the CEO does not ask the board for help (i.e., $m \in \{m_a, m_c\}$), the board is unaware of the deal and, thus, cannot act.[23]

Board's firing decision. At Date (iii) of $t = 1$, the board decides whether to replace the CEO. If the board is "weak" (state of the world ω^{weak}), the CEO is retained. If the board is "tough" (state of the world ω^{tough}), the CEO is fired with probability $\rho(\hat{m}) \in [0, 1]$, where $\hat{m} \in \{m_b, \overline{m}_b\}$ and $\rho(\cdot)$ is optimally chosen by the board to maximize the second period expected profit. CEOs do not want to be replaced because they would lose the nonpecuniary private benefit of $B > 0$ in $t = 2$.

An important timing assumption is that the decision to fire or retain the CEO happens at Date (iii), which is before cash flows are realized (Date (iv)). The interpretation is that corporate policies may have long-term consequences, and decisions to retain a CEO may be made before the full impact of these actions is known. In other words, even if the board could fire the CEO once $y_1 = 0$ is revealed, the CEO would have stayed long enough in the job to enjoy B. Formally, all we need is that the CEO's decision to ask the board for help at Date (ii) has incremental informativeness for the board's decision to fire or retain the CEO. Our timing assumptions retain this property in the simplest possible manner.

---

[23] Even if the board knows that the CEO has identified a deal, it is unaware of x unless the CEO shares this information with the board.

3 AI and Board Learning

In this section, we begin our analysis by focusing on the board's learning problem. In period $t = 1$, the CEO learns more about her own knowledge by trying to solve a problem at Date (ii). If the CEO learns she cannot solve a problem (i.e., her *ex post* type is b), she decides whether to seek help from the AI agent ($m = m_a$) or the board ($m = m_b$). To focus on the more interesting case, we make the following assumption:

Assumption 2 (Advice autonomy threshold). *The board has greater problem-solving capability than the AI agent:* $1 - \eta > \kappa$.

Assumption 2 implies that the AI's problem-solving capability is below of what Garicano, Li, and Wu (2026) call the "autonomy threshold": the point at which full delegation to AI dominates human performance. Under this assumption, a CEO who identifies a deal she cannot assess faces a trade-off. On the one hand, if she asks the board for help, she will receive better advice than the AI agent's. On the other hand, the board may learn that she was unable to assess the deal. In that case, the board may decide to fire the CEO on Date (iii). If Assumption 2 does not hold, no such trade-off exists, and the CEO does not ask the board for help. In contrast, under Assumption 2, the board's advice remains valuable.

The unconditional probability that a randomly chosen CEO is of *ex post* type g is

$$\Pr(g) = \Pr(x \leq \alpha z) = \int_0^1 \int_0^1 \mathbb{1}_{x \leq \alpha z} \, dx \, dz. \tag{1}$$

For a given z, the CEO solves all problems with $x \leq \alpha z$. Thus,

$$\Pr(g) = \int_0^1 \min(\alpha z, 1) \, dz = \int_0^{1/\alpha} \alpha z \, dz + \int_{1/\alpha}^1 1 \, dz = 1 - \frac{1}{2\alpha}. \tag{2}$$

As expected, if the AI agent has greater skill-augmenting capabilities, the probability that the CEO solves the problem is higher. In contrast, and by definition, the AI agent's problem-solving capability does not affect the CEO's ability to solve the problem independently.

Let S_2 denote the event in which a retained CEO can solve the problem in $t = 2$. Let $\Pr(S_2 \mid \tau)$ denote the probability of S_2 conditional on the CEO being "good" or "bad," $\tau \in \{g, b\}$. We have the following result.

Lemma 1 (Learning). *If the CEO from $t = 1$ is retained in $t = 2$, the probability that she can solve a newly drawn problem is $\Pr(S_2|g) = \frac{2(3\alpha-2)}{3(2\alpha-1)}$ if she is good and $\Pr(S_2|b) = \frac{1}{3}$ if she is bad.*

Note that the AI agent's skill-augmenting capability makes the "good" CEO better, but has no effect on the expected output of a "bad" CEO in period 2. The latter result arises because when AI is more skill-augmenting (larger α), the CEO who fails to solve the problem is, in expectation, of even lower type. This negative selection effect exactly cancels out the positive skill-augmenting effect in period 2.

When facing a bad CEO, if the board is "tough," it has the option to replace the CEO with a new one of unknown type. A newly-hired CEO has an average type of $E[z] = \frac{1}{2}$; call this type $\tau = n$. Note that at $t = 2$ there are no career concerns, so CEOs always ask the board for help if they cannot solve a problem. With a newly-hired CEO, the expected output is $E[y_2 \mid n] = (1 - \frac{1}{2\alpha}) + \frac{1}{2\alpha}(1 - \eta) = 1 - \frac{\eta}{2\alpha}$. If a bad incumbent CEO is retained, the expected output is $E[y_2 \mid b] = 1 - \frac{2\eta}{3}$. We define the expected output gain from replacing a bad CEO with a CEO of average type as

$$M(\alpha, \eta) := 1 - \frac{\eta}{2\alpha} - 1 + \frac{2\eta}{3} = \eta\Big(\frac{2}{3} - \frac{1}{2\alpha}\Big) > 0. \tag{3}$$

We can interpret $M(\alpha, \eta)$ as the value of the option to replace a bad CEO, or the *(marginal) benefit of monitoring*. Note that this value increases with the AI agent's skill-augmenting capability, α, and with the cost of board advising, η. By contrast, AI's problem-solving capability does not affect the value of board learning.

To induce a CEO to exert effort in $t = 2$, shareholders must offer her a contract with a sufficiently high output-contingent bonus. To focus on the more interesting case, we make the following assumption.

Assumption 3 (Efficient effort provision). *Exerting effort is always efficient:* $1 - \frac{2}{3}\eta \geq c$.

Assumption 3 implies that it is efficient for a "bad" CEO to exert effort in $t = 2$. That is, shareholders find it optimal to offer a bonus that induces all CEO types to exert effort in $t = 2$.[24]

Under our assumptions, an equilibrium without board learning does not exist, as we show next.

Lemma 2 (No uninformative equilibrium). *There is no equilibrium with positive effort provision in both periods in which* $p_t(b) = p_t(g)$ *for* $t \in \{1, 2\}$.

Recall that $p_t(\tau)$ is the probability that a CEO with *ex post* type $\tau \in \{b, g\}$ asks the board for advice (i.e., $m = m_b$). Lemma 2 implies that CEOs of different *ex post* types must communicate with the board with different probabilities. It then follows that there is no full pooling of types in equilibrium; with some positive probability, the board must learn something about τ from the equilibrium play. In equilibrium, there could be either full separation of types or partial separation, in which case both types may choose the same action, but with different probabilities.

4 The Pre-AI Boardroom

In this section, we consider the case of a "pre-AI economy," where $\alpha = 1$ and $\kappa = 0$. This case serves as a benchmark. Solving this simpler version reveals the mechanisms more clearly and helps build intuition. We extend the benchmark model to the general case in the next section.

---

[24] Nothing important changes if Assumption 3 does not hold, as long as it is efficient to induce the average type (i.e., $z = 0.5$) to exert effort. We impose Assumption 3 simply to reduce the number of cases to consider.

At the beginning of the first period, shareholders choose a pair (w_1, π). For each (w_1, π), players play a sequential game of incomplete information. The solution concept is Perfect Bayesian Equilibrium.

To solve for an equilibrium, note first that because $\kappa = 0$, at Date (ii) of either period action m_c weakly dominates m_a, and thus only two choices are relevant: $m \in \{m_b, m_c\}$. Because there are no career concerns in $t = 2$, the CEO's decision whether to ask the board for help in this period is trivial. If the CEO knows how to solve the problem, she will choose m_c because it delivers w_2 with probability 1, while choosing m_b delivers w_2 with probability $1 - \eta$. Similarly, if the CEO does not know how to solve the problem, it is strictly better to ask for help and gain w_2 with probability $1 - \eta$.

In the first period, at Date (iii), the board updates its beliefs about the CEO after observing $\hat{m} = m \in \{m_b, m_c\}$. Lemma 2 implies that, if effort is provided in the first period (which is the only equilibrium type we need to consider), both nodes at this date (m_b, or "ask for help," and m_c, or "do not ask for help") must be reached with positive probability in equilibrium. Thus, the board's beliefs about the CEO's type are fully pinned down by Bayesian updating, and there are no off-path beliefs to be determined.[25] Because the game is sequential and beliefs are uniquely determined, for each (w_1, π) there is a (generically) unique equilibrium of the continuation game that follows. Since shareholders move first, they select their preferred equilibria by choosing (w_1, π). We refer to the shareholders' preferred equilibrium as the *optimal equilibrium*.[26]

We have the following result.

Lemma 3 (Good CEOs don't ask for help). *In equilibrium,* $p_1(g) = 0$.

Intuitively, because ex post "good" CEOs are better at solving problems than the board,

---

[25] If we changed the timing slightly to allow the firm to determine w_2 for an incumbent CEO after observing y_1, the information node $(m = m_c, y_1 = 0)$ would be off-path for the equilibrium we analyze. Several reasonable off-path beliefs would sustain the equilibrium we study, such as believing that $(m = m_c, y_1 = 0)$ indicates a CEO who did not exert effort in $t = 1$.

[26] Generic uniqueness allows for a measure zero of pairs (w_1, π) to induce multiple equilibria in the continuation game. We will show in the proof of Proposition 1 that this multiplicity does not create difficulties for finding the optimal equilibrium, which is always unique.

not asking the board for help must be a positive signal to the board. By contrast, ex post "bad" CEOs face a trade-off: while asking for help increases the likelihood of solving the problem, it also increases the probability of being fired. Thus, only type-b CEOs can choose mixed strategies in equilibrium: while $p_1(g) = 0$, we must have $p_1(b) > 0$ (due to Lemma 2). It follows immediately from Lemma 3 that after observing m_b, the board knows that the CEO must be of the bad type. Thus, a tough board should always replace a CEO who asks for help: $\rho(m_b) = 1$.

To streamline the exposition, we will henceforth restrict the analysis to pure-strategy equilibria. As we will formally show in the proof of Proposition 1, restricting the analysis to pure strategies is without loss of generality because an optimal equilibrium must be in pure strategies.

Under pure strategies, Lemmas 2 and 3 jointly imply $p_1(b) = 1$, that is, an *ex post* bad CEO always asks the board for help. That is, in equilibrium, bad and good CEOs must make different choices at Date (ii), and these choices fully reveal the CEO's type. Accordingly, we solve the model under the assumption that shareholders expect a bad CEO to request the board's help at $t = 1$. Then, we show that this belief is consistent with equilibrium play, that is, $p_1(b) = 1$.

Solving the model requires checking six incentive compatibility constraints. Working backwards, we begin at Date (ii) in $t = 2$ (note that Date (iii) is redundant in $t = 2$ because the CEO cannot stay for another period). At this point, the CEO has no career concerns and thus does not need to be incentivized to ask the board for help when she is unable to assess a deal. However, the CEO still needs to be incentivized to exert effort c to identify a deal.

Suppose first that the CEO from $t = 1$ is retained in $t = 2$. If the CEO is of type b, from Lemma 1 we know that $E[z \mid b] = \Pr(S_2 \mid b) = \frac{1}{3}$. Let w_{2b} denote the bonus offered to a CEO of type b conditional on $y_2 = 1$. The probability that the output is $y_2 = 1$ if the CEO is of type b is $\frac{1}{3} + \frac{2}{3}(1 - \eta) = 1 - \frac{2}{3}\eta$. A bad CEO's "effort incentive constraint" at Date (ii) of $t = 2$ is

$$\left(1 - \frac{2}{3}\eta\right) w_{2b} - c + B \geq B. \tag{4}$$

The left-hand side of (4) is the bad CEO's payoff in $t = 2$ if she exerts effort. By exerting effort at a personal cost c, the CEO produces $y_2 = 1$ with probability $1 - \frac{2}{3}\eta$, in which case she earns w_{2b}, on top of her private benefit, B. If she does not exert effort, she collects her private benefit, which is the right-hand side of (4). At Date (i) of $t = 2$, if shareholders know that the CEO is of type b, they will offer the lowest possible bonus such that (4) holds, implying $w_{2b}^* = \frac{3c}{3-2\eta}$.

Similarly, if a retained CEO is of type g, from Lemma 1 we know that $E[z \mid g] = \Pr(S_2 \mid g) = \frac{2}{3}$, and the effort incentive constraint is

$$\left(1 - \frac{\eta}{3}\right) w_{2g} - c + B \geq B, \tag{5}$$

implying $w_{2g}^* = \frac{3c}{3-\eta}$.

In the case of a newly appointed CEO, which we denote by type n, we have that $E[z \mid n] = \Pr(S_2 \mid n) = \frac{1}{2}$, and the effort incentive constraint is

$$\left(1 - \frac{\eta}{2}\right) w_{2n} - c + B \geq B, \tag{6}$$

implying $w_{2n}^* = \frac{2c}{2-\eta}$.

By definition, a "weak" board will never replace the CEO. At Date (iii) of $t = 1$, because the marginal benefit of monitoring is positive ($M(1, \eta) > 0$), a "tough" board will always replace a CEO of type b, but will not replace a CEO of type g. Intuitively, there is no reason for a board to retain an *ex post* bad CEO because an outside replacement is expected to be more qualified. Similarly, an *ex post* good CEO is always retained because she is expected to be more qualified than outsiders.[27] That is, in equilibrium we must have $\rho(m_b) = 1$

---

[27] An implicit assumption here is that only the firm (or its board) observes m and the decision of whether to retain the CEO, implying that an *ex post* good CEO does not have better outside opportunities. Allowing CEO retention decisions to signal quality to competing employers (as in Waldman (1984), Greenwald

and $\rho(m_c) = 0$.

Consider now Date (ii) of $t = 1$. Recall that w_1 is the bonus conditional on $y_1 = 1$. First, we need to ensure that a "good" CEO does not want to pretend she is "bad":

$$w_1 + B \geq (1-\eta)w_1 + (1-\pi)\Big[\Big(1 - \frac{\eta}{3}\Big)w^*_{2b} - c + B\Big]. \tag{7}$$

The left-hand side of (7) is the (incremental) payoff a good CEO expects by solving the problem herself, without involving the board. In that case, the CEO guarantees w_1 and knows that she will be retained in $t = 2$. She will then enjoy her private benefit B and exert effort c in return for a second-period bonus of w_{2g}. In that case, her $t = 2$ utility is $\big(1 - \frac{\eta}{3}\big)w^*_{2g} - c + B = B$. The right-hand side of (7) is the payoff a good CEO expects by involving the board. If a good CEO asks the board for help, the board solves the problem with probability $1 - \eta$ and, with probability $1 - \pi$, it retains the CEO (whom the board now thinks is bad) and offers her a second-period bonus of w^*_{2b}.[28]

Second, still on Date (ii), a bad CEO must be incentivized to ask the board for help:

$$(1-\eta)w_1 + (1-\pi)B \geq B. \tag{8}$$

Condition (8) is the "(asking for) help incentive constraint." If a CEO who cannot assess a deal asks the board for help, the board will solve the problem with probability $1 - \eta$, in which case the CEO receives w_1. By revealing her type, the CEO is now retained with probability $1 - \pi$, in which case she enjoys B.[29] If the CEO chooses not to ask for help, she will not solve the problem, but will be retained with probability 1, because $\rho(m_c) = 0$. If retained, the CEO is guaranteed to enjoy her private benefit B at $t = 2$. Because the board

---

(1986), Li (2013), and Ferreira and Nikolowa (2023), among others) complicates the analysis but does not add significant insights.

[28] We assume that when a good CEO asks the board for help, the assessment of the deal is delegated to the board. The board succeeds at assessing the deal with probability $1 - \eta$; the CEO's knowledge here is irrelevant.

[29] The CEO is also promised bonus w^*_{2b} in $t = 2$ in exchange for effort c, but because the effort constraint in (4) binds, the CEO's exactly breaks even.

thinks the CEO is good, it will offer her a bonus of w_{2g}^*. The CEO will choose not to exert effort because her IC in (4) is violated at w_{2g}^*.

The help IC constraint can also be written as

$$\pi \leq \frac{(1-\eta)w_1}{B}. \tag{9}$$

Intuitively, for the CEO to share information with the board, the board must be sufficiently "friendly," that is, its monitoring intensity π should not be too high. The right-hand side of (9) is the ratio between the CEO's benefit from the board's help and the cost of sharing information with the board. A "bad" CEO benefits from asking the board for help because the board can solve a problem with probability $1-\eta$, in which case the CEO receives w_1. The cost of sharing information is that, if the board is "tough," the CEO will be fired and lose her private benefit, B. This logic is reminiscent of Adams and Ferreira's (2007) model.

The sixth and final constraint is the first-period "effort incentive compatibility" constraint:

$$\underbrace{B-c}_{t=1\text{ utility}} + \frac{1}{2}\underbrace{(w_1+B)}_{\text{g's utility}} + \frac{1}{2}\underbrace{((1-\eta)w_1+(1-\pi)B)}_{\text{b's utility}} \geq 2B. \tag{10}$$

The left-hand side of (10) is the CEO's expected lifetime payoff if she exerts effort at $t=1$ and behaves optimally thereafter. By exerting effort, the CEO identifies a deal and eventually learns her type (g or b), each with equal probabilities. The right-hand side of (10) is the CEO's payoff if she chooses instead not to exert effort at $t=1$. In that case, the first-period output is $y_1=0$, implying that she does not earn w_1. In addition, she does not learn her type; she believes her knowledge is average ($E[z]=\frac{1}{2}$). Because she does not ask the board for help, she is retained in $t=2$ and offered a bonus of w_{2g}^*. Since w_{2g}^* violates the second-period IC constraint for type n (see (6)), she does not exert effort at $t=2$. Thus, her lifetime expected utility is $2B$.

The shareholders' problem at Date (i) of $t=1$ is to choose a first-period bonus $w_1 \geq 0$ and a board monitoring intensity (i.e., independence) $\pi \in [0,1]$ to maximize its expected profit, subject to constraints (7), (8) and (10), taking w_{2b}^*, w_{2g}^*, and w_{2n}^* as given. This is a linear programming problem. To solve it, we proceed by initially assuming that (7) is slack. After solving for the optimal values, we then check (and confirm) that (7) is indeed slack at these values. With a slight abuse of terminology, we call a pair (w_1, π) a *governance structure*.[30]

Consider first the case in which $B \leq (1-\eta)2c$. In this case, it can be readily checked that, for all $\pi \in [0,1]$, the effort IC constraint (10) implies the help IC constraint (8). Thus, the optimal board monitoring intensity π^* must be a corner solution that binds constraint (10) only. From (11), we see that the marginal benefit of monitoring a bad CEO is $M(1,\eta) = \frac{\eta}{6}$. Thus, the firm wants to monitor more intensively as long as the marginal benefit exceeds the bad CEO's expected loss from monitoring, which is B. If $M(1,\eta) > B$, the optimal board monitoring intensity is $\pi^* = 1$; if $M(1,\eta) < B$, the optimal board structure is $\pi^* = 0$. Thus, although a friendly board might be optimal in this case, the governance structure is determined not by the CEO's incentives to communicate but rather by the incentives for effort provision, unlike Adams and Ferreira (2007).

The interesting case is, of course, when the effort IC constraint (10) does not imply the help IC constraint (8). To make sure that this happens for any value of η, from now on, we make the following assumption.

Assumption 4 (Help IC constraint matters). *The effort IC constraint does not imply the help IC constraint:* $B \geq 2c$*.*

Figure 1 illustrates constraints (8) and (10) on the (w_1, π) plane, for a set of parameters that satisfy Assumption 4. Note that the help IC constraint passes through the origin and has a smaller slope than the effort IC constraint. The incentive-compatible region is the gray area to the right of both boundaries.

---

[30] Strictly speaking, the governance structure should also include the second-period bonuses: (w_{2b}, w_{2g}, w_{2n}).

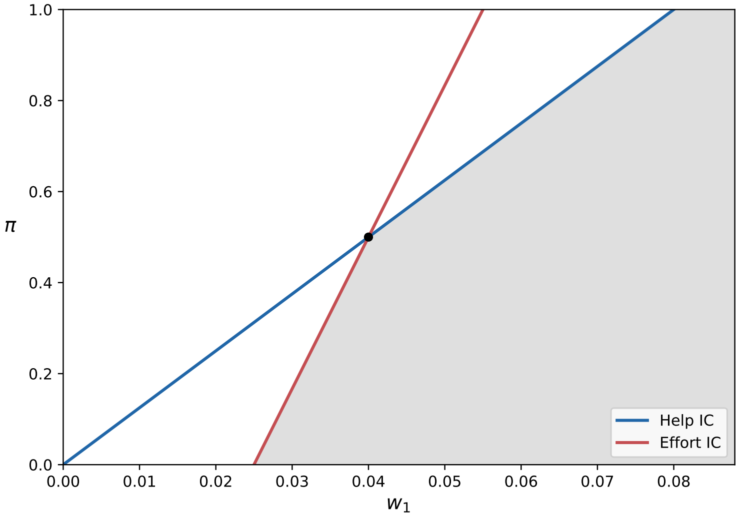

Figure 2: Incentive-compatible governance structures

(Parameters: $\eta = 0.4$, $B = 0.048$, $c = 0.02$)

Given the optimal $t = 2$ bonuses, and assuming that constraints (8) and (10) bind, the firm's expected profit at the beginning of $t = 1$ is

$$\begin{aligned} V(w_1, \pi) &= \frac{1}{2}(\underbrace{2 - \frac{\eta}{3} - c - w_1}_{\text{Profit from good CEO}}) + \frac{1}{2}(\underbrace{2 - \frac{5}{3}\eta - c - (1-\eta)w_1 + \pi\frac{\eta}{6}}_{\text{Profit from bad CEO}}) \\ &= 2 - \eta - c - (1 - \frac{\eta}{2})w_1 + \frac{\pi}{2}M(1, \eta). \end{aligned} \tag{11}$$

From (11) we can write the *isoprofit* for a given level of profit V as

$$\pi(w_1) = \frac{1}{M(1,\eta)}[(2-\eta)w_1 + 2(\eta + c) - 4 + 2V]. \tag{12}$$

The profit increases as we move towards the Northwest in Figure 2. Shareholders must thus choose a governance structure on the highest isoprofit that intersects the incentive-compatible region.

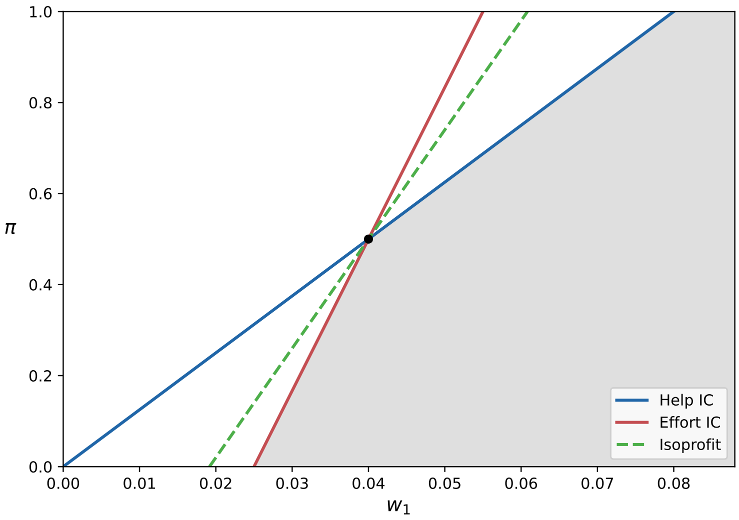

Figure 3: Optimal governance structure when the solution is interior

(Parameters: $\eta = 0.4$, $B = 0.048$, $c = 0.02$)

Figure 3 shows the unique governance structure (w_1^*, π^*) when the solution is interior. This case happens when the isoprofit's slope is between the help IC's slope and the effort IC's slope. This condition can be written as

$$\frac{M(1,\eta)}{B} \in \left(1, \frac{2-\eta}{1-\eta}\right). \tag{13}$$

To understand (13), notice that if the benefit of monitoring $M(1,\eta)$ is less than the cost of monitoring B, then we must be at a corner solution where the optimal monitoring is zero ($\pi^* = 0$). If, instead, $M(1,\eta) > B$, an interior solution obtains unless the benefit of monitoring is "too high" (i.e., $M(1,\eta) > \frac{2-\eta}{1-\eta}B$), in which case the optimal solution requires $\pi^* = 1$.

The following proposition formally characterizes the optimal governance structure.

Proposition 1 (Optimal governance structure). *The (generically unique)*[31] *optimal gover-*

---

[31] In the knife-edge cases of $M(1,\eta) = B$ or $M(1,\eta) = (\frac{2-\eta}{1-\eta})B$, there are multiple values of π that maximize profit.

nance structure is

$$(w_1^*, \pi^*) = \begin{cases} (\frac{2c}{2-\eta}, 0) & \textit{if } M(1,\eta) < B \\ (2c, \frac{(1-\eta)2c}{B}) & \textit{if } \frac{M(1,\eta)}{B} \in (1, \frac{2-\eta}{1-\eta}) \\ (\frac{B}{1-\eta}, 1) & \textit{if } M(1,\eta) > (\frac{2-\eta}{1-\eta})B \end{cases} \tag{14}$$

Proposition 1 shows that a CEO-friendly board ($\pi^* < 1$) is optimal unless the benefit of monitoring is sufficiently high ($M(1,\eta) > \frac{2-\eta}{1-\eta}B$). Under an interior solution, (w_1^*, π^*) solves (8) and (10):

$$w_1^* = 2c \text{ and } \pi^* = \frac{(1-\eta)2c}{B}. \tag{15}$$

Replacing (w_1^*, π^*) in (7) confirms that this constraint is slack (see the Appendix for the proof). The optimal board "friendliness" is $1 - \pi^* > 0$.

5 AI Adoption by the CEO

In this section, we modify the benchmark model to allow the CEO to have access to an AI agent. We start with the simpler case in which the AI agent has problem-solving capabilities ($\kappa > 0$) but no skill-augmenting capabilities ($\alpha = 1$) (Subsection 5.1), before considering the case in which AI has both capabilities (Subsection 5.2). Subsection 5.3 briefly considers an extension in which AI has deal-sourcing capabilities. Since the analysis follows essentially the same steps as in the benchmark case, for brevity, we present only the steps necessary to understand the intuition behind the results.[32]

---

[32] Technically, the AI adoption model introduces a third hidden action, m_a, at Date (ii). Because m_a (m_c) strictly dominates m_c (m_a) when $x > \alpha z$ ($x < \alpha z$), when the board observes $\{m_a, m_c\}$ on the equilibrium path, it must rationally infer that m_c was chosen. Thus, the availability of m_a affects deviation payoffs but nothing else, and all equilibrium selection arguments for the benchmark model apply equally to this case.

5.1 AI as a Coworker

Let $\kappa > 0$ and $\alpha = 1$. The only difference from the benchmark model is that the help IC constraint is now

$$(1-\eta)w_1 + (1-\pi)B \geq \kappa w_1 + B. \tag{16}$$

Intuitively, a CEO who does not know how to assess a deal can now ask the AI agent for help, rather than consulting the board. That is, a bad CEO who chooses not to ask the board for help must choose m_a as it strictly dominates m_c. Conditional on the CEO being bad, the probability that the AI agent solves the problem is κ, thus this option delivers expected payoff $\kappa w_1 + B$, making the help IC constraint harder to satisfy. In turn, shareholders now need a friendlier board to induce the CEO to communicate truthfully with the board.

Figure 4 illustrates the introduction of an AI agent with problem-solving capabilities. The help IC constraint becomes flatter due to $\kappa > 0$, while the effort IC remains the same. The intersection now implies a lower monitoring intensity and a lower first-period bonus in an interior solution. Solving (10) and (16) simultaneously leads to

$$w_{1AI}^* = \frac{2c}{1+\kappa} < w_1^*. \tag{17}$$

and

$$\pi_{AI}^* = \frac{(1-\eta-\kappa)2c}{B(1+\kappa)} < \pi^*. \tag{18}$$

That is, when AI has problem-solving capabilities, in an interior solution, firms must choose a lower monitoring intensity. Intuitively, by offering an imperfect alternative to the board's advice, AI increases the CEO's payoff for not seeking the board's help. This makes the CEO less likely to reveal information that can lead to her dismissal. To encourage the CEO to communicate with the board, the board must become even friendlier. As a consequence, CEOs become more entrenched and are replaced less often. More CEO entrenchment is inefficient because CEOs are kept even when a better replacement is available. The next proposition formally states these results.

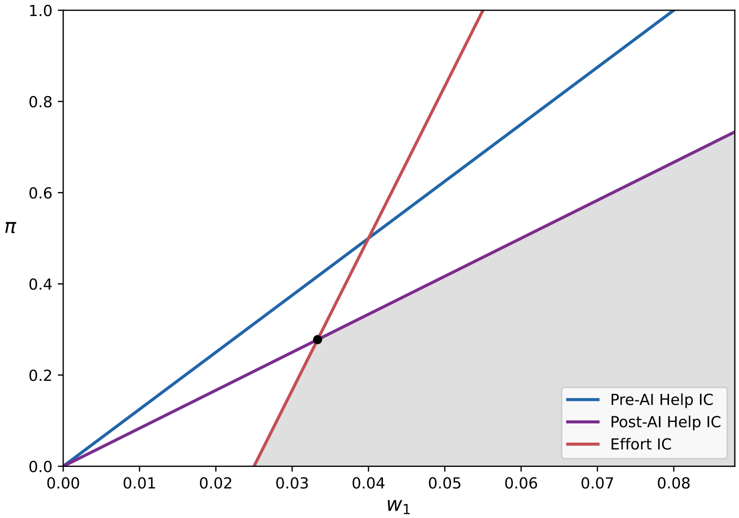

Figure 4: Incentive-compatible region under AI with problem-solving capabilities

(Parameters: $\eta = 0.4$, $B = 0.048$, $c = 0.02$, $\kappa = 0.2$, $\alpha = 1$)

Proposition 2 (Governance and AI's problem-solving capability). *Let* $\kappa \in [0, 1 - \eta]$ *and* $\alpha = 1$. *The optimal governance structure is interior and given by (17) and (18) if and only if*

$$\frac{M(1,\eta)}{B} \in \left(1, \frac{2-\eta}{1-\eta-\kappa}\right). \tag{19}$$

Under an interior solution, an increase in κ *decreases the first-period bonus, the board's monitoring intensity, and the firm's profit. The CEO's expected lifetime utility is independent of* κ.

Proposition 2 implies that improvements in AI's problem-solving capabilities can be inefficient: an increase in κ reduces firm value without affecting the CEO's utility. The firm is worse off with a larger κ because it needs to monitor less intensively to encourage the CEO to communicate with the board. Lower monitoring intensity increases CEO entrenchment, which is inefficient. Because the CEO's lifetime effort IC constraint is binding, the entrenchment cost is fully paid by the firm.

Because the first-period bonus is lower under $\kappa > 0$, while the second-period bonuses are the same, incentive pay falls with κ. We can write the first-period CEO's expected lifetime compensation as

$$E(w) = (1 - \frac{\eta}{2})w^*_{1AI} + \frac{1}{2}(1 - \frac{\eta}{3})w^*_{2g} + \frac{1}{2}(1 - \pi^*_{AI})(1 - \frac{2\eta}{3})w^*_{2b}. \tag{20}$$

The next corollary shows that expected total compensation also falls with κ.

Corollary 1 (**CEO compensation and AI's problem-solving capability**). *Expected total compensation is decreasing in κ.*

Thus, an increase in κ reduces both the average bonus and expected total compensation. That is, while the CEO benefits from increased job stability, her compensation falls as AI's problem-solving capabilities improve. Because the first-period CEO's utility is $2B$ regardless of κ, the CEO is indifferent to the value of κ.

5.2 AI as Both a Tool and a Coworker

We now extend the model to consider the case in which the AI agent possesses skill-augmenting capabilities (i.e., $\alpha \in [1, \overline{\alpha}]$), in addition to its problem-solving capabilities. Note that the help IC constraint in (16) does not change. The second-period bonus for a bad CEO, w^*_{2b}, also does not change. The second-period bonus for a good CEO becomes

$$w^*_{2gAI} = \frac{3c}{3 - \frac{\eta}{2\alpha - 1}} < w^*_{2g} \tag{21}$$

and the second-period bonus for a newly-appointed CEO becomes

$$w^*_{2nAI} = \frac{2c}{2 - \frac{\eta}{\alpha}} < w^*_{2n}. \tag{22}$$

Intuitively, because the CEO is more competent with AI's help, incentivizing her to exert effort is easier, requiring smaller bonuses overall.

The first-period effort incentive compatibility constraint becomes

$$B - c + (1 - \frac{1}{2\alpha})(w_1 + B) + \frac{1}{2\alpha}\left((1-\eta)w_1 + (1-\pi)B\right) \geq 2B. \tag{23}$$

Note that a larger α relaxes the effort IC because it increases the probability that the CEO can solve the problem on her own.

The firm's expected profit at the beginning of $t = 1$ is

$$V(w_1, \pi) = \frac{2\alpha - 1}{2\alpha}\underbrace{\left(2 - \frac{\eta}{3(2\alpha-1)} - c - w_1\right)}_{\text{Profit from good CEO}} + \frac{1}{2\alpha}\underbrace{\left(2 - \frac{5}{3}\eta - c - (1-\eta)w_1 + \pi M(\alpha,\eta)\right)}_{\text{Profit from bad CEO}}. \tag{24}$$

Note that—all else constant—a larger α increases the firm's profit in both states (g and b), and also increases the probability that the CEO is perceived as "good." Because the profit is larger for type g than b for any governance structure (w_1, π), we conclude that α is a positive profit shifter.

The next proposition shows how α affects governance.

Proposition 3 (Governance and AI's skill-augmenting capability). *Let $\kappa \in [0, 1-\eta]$ and $\alpha \geq 1$. The optimal governance structure is interior and given by*

$$w^*_{1AI} = \frac{2c\alpha}{2\alpha - 1 + \kappa} < w^*_1. \tag{25}$$

and

$$\pi^*_{AI} = \frac{(1-\eta-\kappa)2c\alpha}{B(2\alpha - 1 + \kappa)} < \pi^*, \tag{26}$$

if and only if

$$\frac{M(\alpha,\eta)}{B} \in \left(1, \frac{2\alpha - \eta}{1-\eta-\kappa}\right). \tag{27}$$

Under an interior solution, an increase in α decreases the first-period bonus and the board's monitoring intensity, and increases the firm's profit. The CEO's expected lifetime utility is independent of κ.

Proposition 3 shows that the monitoring intensity also decreases with the AI agent's skill-augmenting capability, α. An increase in α makes the CEO more productive, thereby increasing the benefit from exerting effort and thus loosening the CEO's effort IC constraint. This allows the firm to incentivize effort provision with a lower first-period bonus, w_1. But a lower first-period bonus makes it harder to meet the help IC constraint, forcing the firm to decrease board independence, π. Figure 5 illustrates the effect of α on the optimal governance structure.

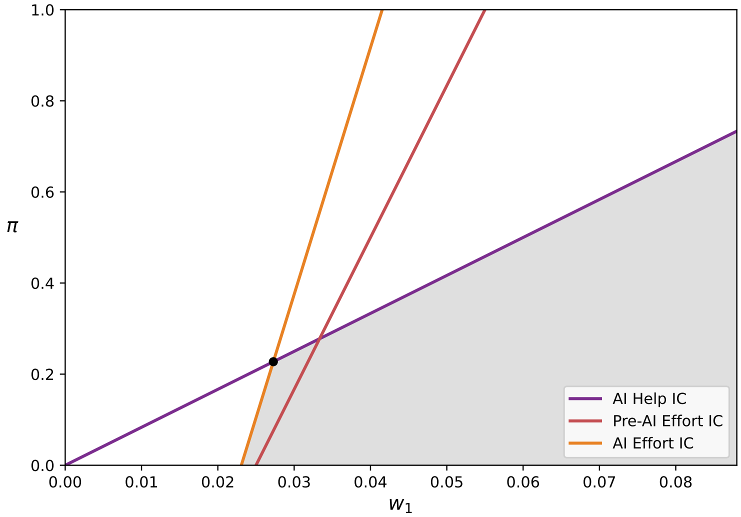

Figure 5: Incentive-compatible region under AI with both capabilities

(Parameters: $\eta = 0.4$, $B = 0.048$, $c = 0.02$, $\kappa = 0.2$, $\alpha = 1.5$)

Overall, as in the case of an increase in κ, both board independence and CEO compensation fall when α increases, as the following corollary shows.

Corollary 2 (**CEO compensation and AI's skill-augmenting capability**). *Expected total compensation decreases with α.*

Despite the equilibrium displaying lower board independence, the firm is better off with a higher α because the increase in CEO productivity, coupled with the fall in compensation (in both periods), more than compensates for the (inefficient) lower CEO turnover rate. This result is robust and not driven by specific parameter assumptions. This occurs because an increase in α both expands the incentive-compatibility region (See Figure 5) and shifts the profit function upward. Intuitively, α expands the technological frontier without (directly) affecting the CEO's incentives to share information. Thus, increasing α must be welfare increasing. Because the effort IC constraint binds, the CEO's utility is unaffected by α, and the firm captures all of the welfare increase as profits.

5.3 AI and Deal Sourcing

Thus far, we have considered the impact of AI on the CEO's ability to assess a deal (i.e., to solve a problem). In our model, the CEO must first exert effort at a cost c to identify a deal. While we assume that deal sourcing cannot be fully automated, it is still reasonable to expect that the CEO's adoption of AI will decrease her cost of deal sourcing.

Suppose that AI's skill-augmenting capability can also be used to reduce the deal-sourcing cost, c. As we can see directly from (17) and (18), a decrease in c decreases both w^*_{1AI} and π^*_{1AI}. Intuitively, a decrease in c has an effect similar to an increase in α: it slackens all effort incentive constraints without affecting the help IC constraint.[33] We conclude that our predictions do not depend on the channel through which AI augments the CEO's skills: if AI enhances the CEO's ability to solve problems or reduces her cost of finding deals, both CEO compensation and board monitoring decrease.

The analysis is more nuanced if we allow an AI agent to substitute for the CEO by independently identifying deals. Suppose that an AI agent can autonomously identify a deal with probability $\phi \in [0, 1]$ at no cost. This case is arguably less realistic, as it conflicts with our assumption that deal sourcing is a "messy job" that needs to be performed by humans. Still, for completeness, we provide a brief analysis of this case.

The immediate effect of allowing the CEO to (covertly) delegate deal sourcing to an AI agent with deal-sourcing capability ϕ is to tighten all incentive constraints. Thus, if the

---

[33] Unlike c, α also affects the probabilities of the CEO being of types b and g, as well as the value of monitoring.

optimal contract requires the CEO to exert effort instead of delegating deal sourcing to the AI agent, an increase in ϕ must increase all bonuses. If the effect of ϕ on w_1 is sufficiently strong, π will also increase with ϕ through the help IC constraint. However, as ϕ increases, inducing the CEO to exert effort must eventually become suboptimal. Intuitively, as ϕ increases, the AI agent becomes an ever cheaper alternative to the CEO for finding deals. At some point, the optimal contract must ignore the effort IC constraints. In that case, both CEO compensation and board monitoring must fall to their minimum possible values. The following proposition formalizes these results.

Proposition 4 (**Governance and AI's deal-sourcing capability**). *Suppose the AI agent can identify a deal with probability* $\phi \in [0, 1]$. *Assume that (13) holds, so that an interior solution exists if* $\phi = 0$. *The following holds:*

1. *If* ϕ *is sufficiently small, both* w_1^* *and* π^* *increase with* ϕ.
2. *There exists* $\overline{\phi} < 1$ *such that for* $\phi \geq \overline{\phi}$, $w_1^* = 0$ *and* $\pi^* = 0$.

Taken together, the results in this subsection show that AI's impact on governance through the deal-sourcing channel depends on whether AI complements or substitutes for the CEO's effort. When AI reduces the CEO's cost of sourcing deals, the effect mirrors that of skill augmentation: both compensation and monitoring decrease, reinforcing our earlier findings. When AI can autonomously identify deals, however, the relationship is non-monotonic. For low levels of autonomous deal-sourcing capability, the need to prevent the CEO from covertly delegating to the AI agent drives both compensation and monitoring upward. Beyond the "autonomy threshold," however, the firm optimally stops incentivizing the CEO to source deals altogether, and both compensation and monitoring collapse to zero (or their minimum possible values). Thus, the long-run effect of AI through deal sourcing is always to reduce both CEO compensation and board monitoring.

5.4 Summary of Empirical Predictions

To summarize, our analysis indicates that improvements in AI (both κ and α) are associated with lower board independence and CEO turnover. AI improvements also reduce CEO incentive compensation, regardless of which capability is enhanced (κ or α). These predictions are reinforced when AI reduces the CEO's cost of sourcing deals (lower c), and they also hold in the long run when AI can autonomously identify deals with probability ϕ—though in the latter case, the short-run effects are non-monotonic, with both compensation and monitoring initially increasing before collapsing.

The welfare implications of AI adoption and improvements are ambiguous, as they depend on the relative strength of each AI capability (κ or α) and how these capabilities are affected by technological advances. If CEOs use AI primarily as a substitute for the board's advice, total welfare is reduced.

From an empirical perspective, we note that these predictions relate to the CEO's adoption of AI, all else held constant. If a shock increases AI usage across the firm as a whole, e.g., knowledge workers use AI more intensively, our predictions may not hold because the resulting improvements in firm productivity may indirectly affect the determination of optimal governance structures. Thus, empirical tests of our predictions must identify shocks to CEOs' AI use that are orthogonal to other shocks affecting firm productivity and governance.

6 AI Adoption by the Board

So far, we have assumed that AI adoption only benefits the CEO. But what if board members can also use AI when performing their tasks? In this section, we extend the model to consider how AI affects the board's performance in its advisory role (Subsection 6.1) and monitoring role (Subsection 6.2).

6.1 AI and the Board's Advisory Role

Suppose the board has access to an AI agent with problem-solving capability κ. If the CEO asks the board for advice, the board can either spend one unit of time assessing the deal x, in which case output is $y_t = 1$ with probability $1 - \eta$, or ask the AI agent for an assessment, in which case output is $y_t = 1$ with probability κ. Assumption 2 ($1 - \eta > \kappa$) implies that the board always prefers to assess the deal rather than delegate this task to the AI agent. Thus, AI's problem-solving capabilities are irrelevant to the board.

Things are more interesting when the board can benefit from AI's skill-augmenting capabilities. Although we have assumed that the board's knowledge is 1, AI can still reduce the cost of board advice by reducing η. For example, AI may reduce the time the board spends on routine compliance tasks. Alternatively, we may assume that the board's knowledge is less than 1, and that the AI agent's skill-augmenting capability enables the board to solve harder problems *and* perform more routine compliance tasks per unit of time.

For simplicity, assume that the board's problem-solving capability is $\beta(1 - \eta)$, where $\beta \in [1, \overline{\beta}]$ is the AI's *board skill-augmenting capability*. To find an interior equilibrium, we replace $1 - \eta$ with $\beta(1 - \eta)$ in the help IC constraint (16) and the effort IC constraint (23). The optimal first-period bonus is unchanged and given by (25). The optimal monitoring intensity becomes

$$\pi^*_{AI} = \frac{(\beta(1 - \eta) - \kappa)2c\alpha}{B(2\alpha - 1 + \kappa)}. \tag{28}$$

It is immediate that increasing β increases the optimal monitoring intensity. Intuitively, an increase in β makes the board a better advisor, providing further incentives for the CEO to ask for help. Thus, all else constant, shareholders can choose a more independent board while still meeting the CEO's help IC constraint. We conclude that board AI adoption can mitigate some of the adverse effects of CEO AI adoption.

Because advances in AI are likely to shift (α, κ, β) in the same direction (but perhaps by different magnitudes), the comparative static exercises we have performed thus far have some apparent limitations. We are interested in shocks that improve (α, κ, β) simultaneously. Let a denote a latent variable measuring AI's overall capabilities. The interpretation is that a higher a increases each of AI's specific capabilities. The effect of changes in a on the monitoring intensity π^*_{AI} depends on the particular correlation structure among the elements in $(\alpha(a), \kappa(a), \beta(a))$. Define the *relative knowledge advantage* the board has over AI as an advisor as

$$R(a) = \beta(a)(1-\eta) - \kappa(a). \tag{29}$$

Assumption 2 and $\beta(a) \in [1, \overline{\beta}]$ imply $R(a) > 0$. We say that a shock $\epsilon_a > 0$ to the latent variable a is *biased against the board* if $R(a+\epsilon_a) - R(a) < 0$, *biased towards the board* if $R(a+\epsilon_a) - R(a) > 0$, and *neutral* if $R(a+\epsilon_a) - R(a) = 0$. Assuming differentiability, as $\epsilon_a \to 0$, the direction of the bias is determined by the sign of $R'(a)$.

We then have

$$\frac{\partial \pi^*_{AI}}{\partial a} = \frac{2c}{B(2\alpha(a) - 1 + \kappa(a))}\Big(R'(a)\alpha(a) - R(a)\frac{\alpha'(a)(1-\kappa(a)) + \kappa'(a)\alpha(a)}{2\alpha(a) - 1 + \kappa(a)}\Big). \tag{30}$$

Note that since $\alpha'(a)(1-\kappa(a)) + \kappa'(a)\alpha(a) > 0$, an AI improvement can increase the board's monitoring intensity only if it is *sufficiently* biased towards the board, that is, if

$$R'(a) > R(a)\frac{\alpha'(a)(1-\kappa(a)) + \kappa'(a)\alpha(a)}{\alpha(a)(2\alpha(a) - 1 + \kappa(a))}. \tag{31}$$

To help with interpretation, we rearrange (31) and simplify notation to get

$$\varepsilon_{Ra} \geq \frac{|\varepsilon_{w_1\alpha}|}{1-\kappa}\Big((1-\kappa)\varepsilon_{\alpha a} + \kappa\varepsilon_{\kappa a}\Big), \tag{32}$$

where ε_{yx} denotes the elasticity of variable y with respect to parameter x. Because the right-hand side of (32) is positive, for π^*_{AI} to increase with advances in AI, we need the board's relative knowledge advantage to be sufficiently responsive to AI technology shocks (i.e., high ε_{Ra}). This condition is harder to obtain when (i) the first-period bonus is very sensitive to AI's skill-augmenting capability (i.e., high $|\varepsilon_{w_1\alpha}|$), (ii) AI's skill-augmenting capability is very sensitive to AI technology shocks (i.e., high $\varepsilon_{\alpha a}$), and (iii) AI's problem-solving capability is very sensitive to AI technology shocks (i.e., high $\varepsilon_{\kappa a}$).

A natural benchmark is to consider changes in a that are *unbiased*, that is, changes in a such that $\varepsilon_{Ra} = 0$. In that case, we have

$$\frac{\partial \pi^*_{AI}}{\partial a} = -2cR(a)\frac{\alpha'(a)(1-\kappa(a)) + \kappa'(a)\alpha(a)}{B(2\alpha(a) - 1 + \kappa(a))^2} < 0. \tag{33}$$

We thus conclude that a technological change that improves AI capabilities in an unbiased manner will lead to a decrease in the board's monitoring intensity. To put it differently, an AI improvement can increase the board's monitoring intensity only if it is sufficiently biased towards the board, that is, if the board's relative advantage over AI as an advisor increases significantly.[34]

To understand the intuition, note that an increase in κ affects the optimal monitoring intensity through two routes. First, it reduces incentives to ask for help for a given first-period bonus. Second, it reduces the optimal first-period bonus, further reducing incentives to ask for help. In contrast, an increase in β increases incentives to ask for help for a given bonus, but does not affect the optimal bonus (note that the optimal bonus (25) is independent of η as well). Therefore, a shock of similar magnitude to both κ and $\beta(1-\eta)$ would not change incentives to ask for help for a given first-period bonus, but would still lower the bonus, tightening the help IC constraint. The net effect on the monitoring intensity is thus negative. In addition, if the same shock also increases α, the first-period bonus is further reduced, further contributing to the tightening of the help IC constraint.

---

[34] A natural assumption is that technological improvements measured by a have decreasing marginal returns on skill-augmentation (for both the CEO and the board). Under that assumption, because $\beta(1-\eta) > \kappa$, an increase in a benefits the CEO more than the board, implying $R'(a) < 0$.

6.2 AI and the Board's Monitoring Role

We now consider the case in which AI can assist the board in performing its monitoring task. Suppose that, with the help of AI, the board may independently learn some of the information the CEO possesses. For example, with AI, the board may find it easier to process market information about the firm.[35] We maintain the assumption that the CEO must exert effort to identify a deal. That is, the board cannot identify a deal unknown to the CEO. For concreteness, suppose that at some point before Date (iii), with probability $\lambda \in [0,1]$ the board learns both about the existence of a deal (with unknown difficulty x) and whether $z \geq x$ or $z < x$, that is, whether the CEO can assess the deal.[36] At first blush, it may seem that a bad CEO would now be more willing to ask the board for help, because the board may learn that the CEO is bad even without the CEO asking for help. However, this reasoning is incomplete because when the board learns about a problem the CEO cannot solve, the board can solve it with probability $\beta(1-\eta)$, which benefits the CEO through her first-period bonus w_1.

The help IC constraint is

$$\beta(1-\eta)w_1 + (1-\pi)B \geq (1-\lambda)(\kappa w_1 + B) + \lambda\left(\beta(1-\eta)w_1 + (1-\pi)B\right), \tag{34}$$

which implies

$$\beta(1-\eta)w_1 + (1-\pi)B \geq \kappa w_1 + B. \tag{35}$$

That is, the board's ability to independently obtain the CEO's information has no impact on the help IC constraint. In an equilibrium in which the help IC constraint is met, the CEO always shares information with the board, thus λ also does not affect the effort IC constraints or the isoprofits, implying that the equilibrium is unaffected by λ.

---

[35] See Ferreira, Ferreira, and Raposo (2011) for evidence that the board's monitoring function is helped by information embedded in stock prices.

[36] It is less reasonable to assume that the board can learn about the CEO's ability to assess a particular deal without learning also what the deal is.

The intuition behind this result stems from the fact that the help IC constraint must hold in the optimal contract. When this constraint holds, the CEO's expected payoff from seeking the board's help is at least as large as from withholding information. Consequently, the CEO weakly prefers full disclosure to probabilistic disclosure. We can interpret λ as the likelihood that the board discovers the problem through independent inspection. Crucially, because the CEO's payoff remains unchanged whether the board learns through voluntary disclosure or independent discovery, the help IC constraint continues to hold even when the board can inspect. The optimal contract is independent of the board's inspection capability as long as it induces the CEO to disclose the information voluntarily.

It's worth noting that if the board could obtain verifiable evidence of the CEO's information withholding and could write contracts contingent on such evidence, this would make it easier for the CEO to communicate the information. However, such contractibility would violate a fundamental premise of our model: that the problems CEOs face are inherently non-contractible due to their complexity and the difficulty of ex-ante specification. That is, these are the types of strategic decisions and opportunities that cannot be fully anticipated or described in formal contracts.

Although AI may be helpful to the board in other ways, we conclude that, within the model's logic, the board's ability to obtain information independently of the CEO does not affect the CEO's incentives to exert effort or communicate with the board.

7 Discussion and conclusions

This paper develops a framework for understanding how AI transforms the relationship between CEOs and corporate boards. Such a framework serves as an analytical tool for studying the corporate governance consequences of CEO AI adoption, which is still in its early stages but growing rapidly (c.f. Yotzov et al. (2026)).

Our model captures two distinct technological capabilities of AI: its problem-solving capability, which enables AI to substitute for human judgment, and its skill-augmenting capability, which enhances human decision-making. That is, we model AI as both a tool and as a coworker. We show that both capabilities reduce board monitoring intensity and CEO compensation, but through different mechanisms. While AI's problem-solving capability unambiguously destroys value by making it more difficult for the board to identify and replace bad CEOs, its skill-augmenting capability can enhance firm value despite reducing board monitoring. When boards also adopt AI, some negative effects are mitigated, but only if technological advances are sufficiently biased toward enhancing board capabilities relative to those of the CEO.

Our analysis reveals a fundamental tension in AI adoption: technology that enhances CEOs' capabilities also reduces their dependence on board advice, disrupting traditional information flows that underpin effective governance. Firms respond by reducing board independence to restore communication incentives, thereby increasing CEO entrenchment without corresponding benefits for CEOs. This result adds a cautionary note to the conventional wisdom that technological advancement improves organizational efficiency. Instead, we show that AI adoption necessitates a reconfiguration of governance structures, adjusting both the intensity of monitoring and compensation packages.

The framework we develop extends beyond the CEO-board relationship to any setting where principals need agents to both exert effort and communicate truthfully. Consider a sales manager overseeing field agents who must both pursue leads (effort) and report difficulties in dealing with potential customers (communication). If AI tools enable agents to handle customer interactions more effectively, agents may stop sharing crucial information with managers. Similarly, procurement officers need to both identify suppliers (effort) and evaluate them. Better AI might prompt the procurement officer to bypass marginal suppliers and go directly to the manager, thereby creating operational risks. In each case, our model predicts that organizations must reduce monitoring intensity and adjust compensation to maintain information flows.

A broad theoretical insight from our analysis is that AI can paradoxically increase transaction costs by tightening incentive compatibility constraints. This occurs because AI affects not only the equilibrium payoff of the agent, but also her *off-equilibrium payoff*: the value of not asking for help in our setting. Under our maintained assumption that the board's advice remains superior, the CEO goes to the board on the equilibrium path, so AI functions as an off-equilibrium threat. The governance distortion arises from AI's availability, not its use. Organizations that fail to recognize this and adapt their governance structures accordingly may find that AI adoption leads to worse outcomes than the pre-AI era.

Appendix: Proofs

Proof of Lemma 1: Using the law of total probability, we have

$$\Pr(S_2|g) = \int_0^1 \Pr(S_2|z,g) \cdot f(z|g)\, dz.$$

Since problems are independent given z, then

$$\Pr(S_2|z,g) = \Pr(S_2|z) = \min(\alpha z, 1).$$

From Bayes' theorem, the posterior distribution of z given g is

$$f(z|g) = \frac{\Pr(g|z) \cdot f(z)}{\Pr(g)} = \frac{\min(\alpha z, 1)}{1 - \frac{1}{2\alpha}}.$$

Therefore

$$\Pr(S_2|g) = \frac{1}{1-\frac{1}{2\alpha}} \int_0^1 [\min(\alpha z, 1)]^2\, dz = \frac{1}{1-\frac{1}{2\alpha}} \int_0^{1/\alpha} (\alpha z)^2\, dz + \int_{1/\alpha}^1 1\, dz = \frac{2(3\alpha-2)}{3(2\alpha-1)}.$$

Similarly, we have

$$\Pr(S_2|b) = \int_0^1 \Pr(S_2|z,b) \cdot f(z|b)\, dz.$$

Since problems are independent given z, we have

$$\Pr(S_2|z,b) = \Pr(S_2|z) = \min(\alpha z, 1).$$

If $z > 1/\alpha$, the CEO can solve any problem. Therefore, $f(z|b) = 0$ for $z > 1/\alpha$. For $z \leq 1/\alpha$, we have $f(z|b) = 2\alpha(1-\alpha z)$. Then

$$\Pr(S_2|b) = \int_0^{1/\alpha} \alpha z \cdot 2\alpha(1-\alpha z)\, dz = 2\alpha^2 \left[\frac{z^2}{2} - \frac{\alpha z^3}{3}\right]_0^{1/\alpha} = \frac{1}{3}.$$

□

Proof of Lemma 2: Suppose the board is tough. At Date (iii) of $t = 1$, it must then decide whether to retain or fire the CEO. Suppose $p_1(\tau)$ is uninformative about τ, i.e., $p_1(\tau) = p$ is a constant. Thus, the board expects the incumbent CEO to be good with probability $1 - \frac{1}{2\alpha}$ regardless of whether the board observes m_b or $\{m_a, m_c\}$. Note that, in period 2, a CEO has no career concerns and thus does not need to be incentivized to ask the board for help when she cannot assess a deal. However, the CEO still needs to be incentivized to exert effort c to identify a deal. If a CEO of type τ is to exert effort, her second-period *effort incentive constraint* must hold:

$$\Big(\Pr(S_2 \mid \tau) + (1 - \Pr(S_2 \mid \tau))(1-\eta)\Big) w_2 - c + B \geq B. \tag{A.1}$$

Suppose first that the board fires the CEO and hires a replacement. We denote the type of the replacement by n. We have $\Pr(S_2 \mid n) = 1 - \frac{1}{2\alpha}$. The board will make the new CEO's incentive constraint bind, which implies a second-period expected profit of

$$v_{2n} = 1 - \frac{\eta}{2\alpha} - c. \tag{A.2}$$

Suppose instead that the firm retains the incumbent CEO, and offers a wage that is sufficient to incentivize the bad CEO, $w_{2b} = \frac{3c}{3-2\eta}$ (here we consider the case in which the CEO exerts effort in the first period, so she privately learns about her type). Under this contract, both CEO types, b and g, will exert effort in period 2. Let v_{2gb} denote the second-period profit when the firm retains an incumbent of unknown type and both types are induced to exert effort. After algebra, we get

$$v_{2gb} = \frac{(3-2\eta-3c)(2\alpha-\eta)}{2\alpha(3-2\eta)}$$

and thus

$$v_{2n} - v_{2gb} = \frac{c\eta(4\alpha-3)}{2\alpha(3-2\eta)} > 0,$$

implying that the board strictly prefers hiring a new CEO to retaining a CEO of unknown type with a contract that induces both types to exert effort.

Alternatively, the board may choose w_{2g} such that it incentivizes only the good CEO. In this case, we have expected profit

$$v_{2g} = \frac{2\alpha-1}{2\alpha}(1 - \frac{\eta}{3(2\alpha-1)} - c),$$

and thus

$$v_{2n} - v_{2g} = \frac{1}{2\alpha}(1 - \frac{2\eta}{3} - c) > 0,$$

which follows from Assumption 3. Thus, if the board learns nothing, it always prefers to replace the CEO. But if a bad CEO knows she will be replaced, she will ask for advice, because her expected payoff will be $(1-\eta)w_1 > 0$ (because w_1 must be positive if the CEO is to exert effort in period 1, knowing she will be replaced in period 2). This implies that a bad CEO always shares information with the board, i.e., $p_1(b) = 1$. By contrast, a good CEO prefers not to ask the board for help, because she can solve the problem with probability 1, while the board solves it only with probability $1-\eta$, implying $p_1(g) = 0$, contradicting $p_1(b) = p_1(g) = p$. We conclude that $p_1(g) \neq p_1(b)$.

Because there are no career concerns in $t = 2$, trivially $p_2(b) = 1$ and $p_2(g) = 0$. $\square$

Proof of Lemma 3. Conditional on knowing that her own ex post type is g at $t = 1$, a CEO who chooses m_c expects to receive w_1 with probability 1. By choosing m_b instead, the CEO expects $(1 - \eta)w_1$. Thus, a good CEO would only choose w_b with strictly positive probability if asking the board for help increases her probability of retention, i.e., if $\rho(m_b) < \rho(m_c)$. But if $\rho(m_b) < \rho(m_c)$, action $m = m_b$ is strictly dominant for a CEO of type b, implying $p_1(b) = 1$. Lemma 2 then implies $p_1(g) < 1$. This means that when the board observes m_b, it must revise down its expectation of the CEO's knowledge, as the CEO is more likely to be bad than good. But then choosing $\rho(m_b) < \rho(m_c)$ is not rational for shareholders, because a replacement CEO is equally likely to be good or bad, and the proof of Lemma 2 shows that replacement always dominates retention when the incumbent is equally likely to be good or bad. We conclude that $p_1(g) = 0$. $\square$

Proof of Proposition 1: **Case 1:** $M(1, \eta) < B$. In this case, the isoprofit is steeper than the effort IC constraint. As argued in the text, the optimal π must be zero in this case. This implies that the help IC constraint is slack; the "bad" CEO always asks for help. The optimal first-period bonus is determined by the effort IC constraint alone:

$$B - c + \frac{1}{2}(w_1^* + B) + \frac{1}{2}\left((1 - \eta)w_1^* + B\right) = 2B, \tag{A.3}$$

which implies $w_1^* = \frac{2c}{2-\eta}$. Replacing it in (11) yields $V(w_1^*, 0) = 2 - \eta - 2c$, which is positive given Assumption 3. Because violating the effort IC constraint leads to zero profit, choosing $w_1^* = \frac{2c}{2-\eta}$ and $\pi^* = 0$ must be the unique optimal solution for this set of parameters.

Case 2: $\frac{M(1,\eta)}{B} \in (1, \frac{2-\eta}{1-\eta})$. As illustrated in Figure 3, the tangency between the highest isoprofit and the IC region happens when both (8) and (10) bind. Solving for (w_1^*, π^*) yields:

$$w_1^* = 2c \text{ and } \pi^* = \frac{(1 - \eta)2c}{B}. \tag{A.4}$$

To confirm that constraint (7) is slack, we replace (w_1^*, π^*) in (7) to get slack

$$\frac{c\big(6B - 5B\eta + 2c\eta(1-\eta)\big)}{B(3-2\eta)} > 0 \quad \text{for all } \eta \in (0,1),\ B > 0,\ c > 0.$$

Thus, the IC constraint in (7) is slack at the optimal solution.

To confirm that (A.4) is the unique optimal governance structure in this case, we also need to consider values of (w_1, π) that violate (8) under pure strategies. If any equilibrium also violates the first-period effort IC constraint, its overall profit is bounded above by v_{2n}, which is less than the profit under the optimal pure-strategy equilibrium. Thus, we do not need to consider these "shirking equilibria," as they are strictly dominated.

If (8) is violated, there is no pure-strategy equilibrium (without shirking): with $\rho(m_c) = 0$ the bad type strictly prefers m_c, leading to pooling (ruled out by Lemma 2). With $\rho(m_c) = 1$, m_b is a dominant action for type b, so we must have $p_1(g) < 1$. Then, m_c would reveal type g and $\rho(m_c) = 1$ cannot be sequentially rational. Thus, for such values, equilibria with first-period effort must involve mixed strategies. We now show that such mixed-strategy equilibria always deliver lower profits than some pure-strategy equilibria associated with governance structure (w_1^*, π^*). That is, mixed-strategy equilibria are never optimal equilibria.

Suppose (w_1, π) is such that only a strictly mixed-strategy equilibrium exists where the first-period effort constraint is satisfied. Let $p_1(b) = p \in (0,1)$ denote the equilibrium probability that a CEO of type b asks the board for help. Let $\rho(m_c) = \rho \in (0,1)$ be the equilibrium probability that a CEO who does not ask for help is replaced, conditional on the board being tough ($\omega = \omega^{tough}$). The bad CEO must be indifferent between m_b and m_c:

$$(1-\eta)w_1 + (1-\pi)B = \pi(1-\rho)B + (1-\pi)B, \tag{A.5}$$

which implies $\pi = \frac{(1-\eta)w_1}{(1-\rho)B}$. We now consider two cases.

Suppose first that $w_1 \geq 2c$. If a strict mixed-strategy equilibrium exists for governance structure (w_1, π), a tough board must be indifferent between firing or retaining a CEO who did not ask for help, implying the expected profit from retaining the incumbent must equal $v_{2n} = 1 - \frac{\eta}{2} - c$, which is the expected second-period profit when a new CEO is hired. The weak board after m_c always retains the incumbent, faces the same posterior, and offers the same contract. So its profit from retention is also v_{2n}. Thus, the unconditional expected second-period profit must be

$$v_{2n} - \frac{p}{2}(1-\pi)M(1,\eta).$$

We obtain this expression by noting that the profit after m_c must be v_{2n}, while after m_b it is v_{2n} if the CEO is replaced and v_{2b} if retained. The unconditional probability of retaining a bad CEO is $\frac{p}{2}(1-\pi)$, in which case the firm "loses" the marginal benefit of monitoring, $M(1,\eta)$.

By contrast, under (w_1^*, π^*), the optimal pure-strategy equilibrium implies an expected period 2 profit of

$$\begin{aligned}\frac{1}{2}v_g + \frac{1}{2}(\pi^* v_{2n} + (1-\pi^*)v_{2b}) &= \frac{1}{2}(v_{2n} + M(1,\eta) + \pi^* v_{2n} + (1-\pi^*)(v_{2n} - M(1,\eta)) \\ &= v_{2n} + \frac{\pi^*}{2}M(1,\eta).\end{aligned}$$

We conclude that the second-period profit is strictly larger under the pure-strategy equilibrium for the optimal governance structure, (w_1^*, π^*). Now we need to compare first-period profits. For a mixed-strategy equilibrium under (w_1, π), we have the first-period profit

$$\frac{1}{2}(1-w_1) + \frac{1}{2}p(1-\eta)(1-w_1),$$

while for the optimal pure strategy equilibrium, we have

$$\frac{1}{2}(1-w_1^*) + \frac{1}{2}(1-\eta)(1-w_1^*).$$

Since $w_1^* = 2c$ and $w_1 \geq 2c$, it follows that the first-period profit is also larger under the pure-strategy equilibrium for the optimal governance structure. We conclude that, if $w_1 \geq 2c$, any mixed-strategy equilibrium is dominated by the optimal pure-strategy equilibrium under (w_1^*, π^*).

Suppose now that $w_1 < 2c$. To prove our result, we first find an upper bound to the profit under (w_1, π) and then show that this upper bound is strictly smaller than the profit under the pure-strategy equilibrium induced by (w_1^*, π^*).

Define the *first-best surplus*, S^{FB}, as the total surplus to all players if the game is played cooperatively: (i) the CEO exerts effort, (ii) a bad CEO always asks for help, and (iii) the board always replaces a bad CEO and retains a good CEO. S^{FB} denotes social surplus—output net of effort costs plus CEOs' private benefits. Since the firm employs one CEO each period and the employed CEO enjoys B, at least $2B$ of social surplus is non-extractable by shareholders. Let $\overline{V}^{FB} := S^{FB} - 2B$ denote the *extractable first-best surplus*: the maximum surplus shareholders can appropriate. The logic is that an employed CEO can get at least B in that period by not exerting effort, so at least $2B$ of the surplus must go to CEOs.

In a pure-strategy equilibrium where both (8) and (10) bind, the $t = 1$ expected utility of an employed CEO is $2B$, which is the minimum that a CEO must get to exert effort at $t = 1$. Because (w_1^*, π^*) implies a pure strategy equilibrium where both (8) and (10) bind, the profit is $V(w_1^*, \pi^*) = \overline{V}^{FB} - \frac{1}{2}(1 - \pi^*)M(1, \eta)$. The intuition is that the only source of inefficiency under a pure-strategy optimal equilibrium is that with probability $\frac{1}{2}(1 - \pi^*)$ a bad CEO is retained, which implies a surplus loss of $M(1, \eta)$. In addition, in this equilibrium, shareholders capture all of the extractable surplus.

Define $\overline{V}^m(w_1, \pi)$ as the extractable surplus under a mixed-strategy equilibrium associated with (w_1, π). By definition, the extractable surplus is an upper bound for the profit under this equilibrium: $\overline{V}^m(w_1, \pi) \geq V(w_1, \pi)$.

We have

$$\overline{V}^m(w_1, \pi) = \overline{V}^{FB} - \frac{1 - \pi(p + (1-p)\rho)}{2} M(1, \eta) - \frac{\pi\rho}{2} M(1, \eta) - \frac{1-p}{2}(1 - \eta).$$

We derive this expression by subtracting from the extractable first-best surplus the expected losses from inefficient actions. There are three sources of inefficiency: (i) sometimes a bad CEO is retained (the second term on the right-hand side), (ii) sometimes a good CEO is replaced (the third term on the right-hand side), and (iii) sometimes a bad CEO does not ask for help (the fourth term on the right-hand side). This expression simplifies to

$$\overline{V}^{m}(w_1,\pi)=\overline{V}^{FB}-\frac{1-p\pi(1-\rho)}{2}M(1,\eta)-\frac{1-p}{2}(1-\eta).$$

Now, use the fact that $\pi=\frac{(1-\eta)w_1}{(1-\rho)B}$ and $w_1<2c=w_1^*$ to conclude that $\pi(1-\rho)<\pi^*$, so that

$$\frac{1-p\pi(1-\rho)}{2}>\frac{1-\pi^*}{2},$$

implying $V(w_1^*,\pi^*)>\overline{V}^{m}\geq V(w_1,\pi)$. We conclude that the pure-strategy equilibrium under (w_1^*,π^*) has strictly higher profit than any mixed-strategy equilibrium.

As a final point, note that at (w_1^*,π^*), mixed strategy equilibria may also exist, with $p\in(0,1)$ and $\rho=0$. However, any such equilibrium trivially has a lower profit than the pure-strategy equilibrium. Furthermore, such equilibria are not robust in the following sense. If under (w_1^*,π^*) the CEO is expected to play a strict mixed strategy, then shareholders can choose instead contract $(w_1^*+\varepsilon,\pi^*)$ with $\varepsilon>0$ infinitesimally small, which breaks the bad CEO's indifference between asking or not asking for help, leading to a unique equilibrium in pure strategies where the help IC constraint is (slightly) slack. The profit in this equilibrium is arbitrarily close to $V(w_1^*,\pi^*)$. We thus conclude that such mixed-strategy equilibria cannot be optimal and would never be chosen by shareholders.

Case 3: $M(1,\eta)>\frac{2-\eta}{1-\eta}B$. In this case, the isoprofit is flatter than the help IC constraint and, as argued in the text, the optimal board structure must be $\pi^*=1$. The effort IC constraint is not binding, thus the help IC constraint determines the optimal first-period bonus: $w_1^*=\frac{B}{1-\eta}$. Replacing it in (11) yields $V(\frac{B}{1-\eta},1)=2-\eta-c-\frac{1-\frac{\eta}{2}}{1-\eta}B+\frac{\eta}{12}$. Because $M(1,\eta)>\frac{2-\eta}{1-\eta}B\Rightarrow\frac{\eta}{6}>\frac{2-\eta}{1-\eta}B\Rightarrow\frac{\eta}{12}>\frac{1-\frac{\eta}{2}}{1-\eta}B$, we have $V(\frac{B}{1-\eta},1)>0$. $\square$

Proof of Proposition 2: The effort IC constraint is independent of κ, and its slope is $\frac{2-\eta}{B}$. The isoprofit is also independent of κ, and its slope is $\frac{2-\eta}{M(1,\eta)}$. As we can see from Figure 3, an interior solution requires $M(1,\eta) > B$.

The help IC constraint's slope is now $\frac{1-\eta-\kappa}{B}$. The slope must be smaller than that of the isoprofit, implying

$$\frac{M(1,\eta)}{B} < \frac{2-\eta}{1-\eta-\kappa}.$$

All other steps are similar to those in the proof of Proposition 1 and are omitted for brevity. From (17) we have

$$\frac{\partial w^*_{1AI}}{\partial \kappa} = -\frac{2c}{(1+\kappa)^2} < 0,$$

and from (18) we have

$$\frac{\partial \pi^*_{AI}}{\partial \kappa} = \frac{2c(\eta-2)}{B(1+\kappa)^2} < 0.$$

From (11) we have

$$\frac{\partial V^*_{AI}}{\partial \kappa} = (1-\frac{\eta}{2})\frac{2c}{(1+\kappa)^2} + \frac{c(\eta-2)M(1,\eta)}{B(1+\kappa)^2} = \frac{c(2-\eta)}{(1+\kappa)^2}\Big(1-\frac{M(1,\eta)}{B}\Big) < 0,$$

which follows because $M(1,\eta) > B$ in an interior solution. Because the effort IC constraint in (10) binds, the CEO's lifetime utility is $2B$. □

Proof of Corollary 1: Because the first-period effort IC constraint binds, we must have

$$E(w) + B - c + \frac{1}{2}(B-c) + \frac{1}{2}(1-\pi^*_{AI})(B-c) = 2B,$$

which implies

$$E(w) = 2c + \frac{\pi^*_{AI}}{2}(B-c).$$

Since $B > c$ and π^*_{AI} is decreasing in κ, then $E(w)$ is decreasing in κ. □

Proof of Proposition 3: The effort IC constraint can be written as

$$\pi \leq \frac{(2\alpha - \eta)w_1 - 2\alpha c}{B}. \tag{A.6}$$

The isoprofit now becomes

$$\pi(w_1) = \frac{1}{M(\alpha,\eta)}[(2\alpha - \eta)w_1 + 2(\eta + c\alpha) - 4\alpha + 2\alpha V]. \tag{A.7}$$

As we can see from Figure 3, an interior solution requires the effort IC constraint to be steeper than the isoprofit, which implies $M(\alpha,\eta) > B$.

The help IC constraint is independent of α, and its slope is $\frac{1-\eta-\kappa}{B}$. The slope must be smaller than that of the isoprofit, implying

$$\frac{M(\alpha,\eta)}{B} < \frac{2\alpha - \eta}{1 - \eta - \kappa}.$$

All other steps are similar to those in the proof of Proposition 1 and are omitted for brevity. From (25) we have

$$\frac{\partial w^*_{1AI}}{\partial \alpha} = \frac{2c(\kappa - 1)}{(2\alpha - 1 + \kappa)^2} < 0,$$

and from (26) we have

$$\frac{\partial \pi^*_{AI}}{\partial \alpha} = \frac{(1 - \eta - \kappa)2c(\kappa - 1)}{B(2\alpha - 1 + \kappa)^2} < 0.$$

It can be shown that $\frac{\partial V^*_{AI}}{\partial \alpha} > 0$ by brute force, but that route is long and not illuminating. A simpler and more intuitive (and elegant) proof is to note that increasing α (i) relaxes the effort IC constraint, (ii) does not affect the help IC constraint, and (iii) (strictly) increases profit for any given (w_1, π) (the arguments are presented in text, and follow directly from inspection of these expressions). It thus follows from (i)+(ii)+(iii) that the firm must do strictly better when maximizing profit under a larger α.

Because the effort IC constraint in (23) binds, the CEO's lifetime utility is $2B$. □

Proof of Corollary 2: Because the first-period effort IC constraint binds, we must have

$$E(w) + B - c + (1 - \frac{1}{2\alpha})(B - c) + \frac{1}{2\alpha}(1 - \pi^*_{AI})(B - c) = 2B,$$

which implies

$$E(w) = 2c + \frac{\pi^*_{AI}}{2\alpha}(B - c) = 2c + \frac{(1 - \eta - \kappa)c}{B(2\alpha - 1 + \kappa)}(B - c),$$

which is decreasing in α. □

Proof of Proposition 4. For notational simplicity, we work with the case in which $\alpha = 1$ and $\kappa = 0$; the results are identical in the general case.[37] Suppose the CEO can ask an AI agent to identify a deal. If asked, the agent identifies a deal with probability $\phi \in (0, 1)$. Suppose that (13) holds, so we have an interior solution if $\phi = 0$. Suppose it is optimal to induce the CEO to exert effort under $\phi > 0$, regardless of type. A bad CEO's effort incentive constraint at Date (ii) of $t = 2$ is now

$$\Big(1 - \frac{2}{3}\eta\Big) w_{2b} - c + B \geq \phi\Big(1 - \frac{2}{3}\eta\Big) w_{2b} + B, \tag{A.8}$$

implying $w^*_{2b} = \frac{3c}{(1-\phi)(3-2\eta)}$. Similarly, we also have $w^*_{2g} = \frac{3c}{(1-\phi)(3-\eta)}$ and $w^*_{2n} = \frac{2c}{(1-\phi)(2-\eta)}$.

The help IC constraint is now

$$(1 - \eta)w_1 + (1 - \pi)\Big(\phi\Big(1 - \frac{2}{3}\eta\Big) w^*_{2b} + B\Big) \geq \phi\Big(1 - \frac{2}{3}\eta\Big) w^*_{2g} + B, \tag{A.9}$$

which simplifies to

$$\pi \leq \frac{(1 - \eta)w_1 + \frac{\eta}{3-\eta}\frac{\phi}{1-\phi}c}{\frac{\phi}{1-\phi}c + B}. \tag{A.10}$$

---

[37] We note that the argument in Footnote 18 does not apply to this extension. If we change the timing so that second-period bonus of a retained CEO is determined after observing y_1, the IC constraints become more complex. However, the qualitative conclusions are unchanged.

The first-period effort IC constraint (when the help IC constraint binds) is

$$-c+\frac{1}{2}\Big(w_1+\frac{\phi}{1-\phi}c+B\Big)+\frac{1}{2}\Big(\frac{1-\frac{2\eta}{3}}{1-\frac{\eta}{3}}\frac{\phi}{1-\phi}c+B\Big)=-c+\frac{w_1}{2}+\frac{\phi}{1-\phi}\frac{1-\frac{\eta}{2}}{1-\frac{\eta}{3}}c+B\geq$$

$$\phi(\frac{w_1}{2}+\frac{\phi}{1-\phi}\frac{1-\frac{\eta}{2}}{1-\frac{\eta}{3}}c)+(1-\phi)\Big(\frac{\phi}{1-\phi}\frac{1-\frac{\eta}{2}}{1-\frac{\eta}{3}}c\Big)+B=\phi\frac{w_1}{2}+\frac{\phi}{1-\phi}\frac{1-\frac{\eta}{2}}{1-\frac{\eta}{3}}c+B.$$

To understand the last term, note that if the CEO does not exert effort in period 1, with probability $1-\phi$ she will not know her type in period 2, in which case she will not exert effort in that period as well. In that state, she will find a deal with probability ϕ (through AI), which she can assess with probability $1-\frac{\eta}{2}$, in which case she receives a bonus $w_{2g}^*=\frac{3c}{(1-\phi)(3-\eta)}$ (because the board now thinks the CEO is of type g given the lack of communication).

Solving for w_1^* yields

$$w_1^*=\frac{2c}{1-\phi}.$$

Notice that the bonus is increasing in ϕ; when $\phi=0$, we have $w_1^*=2c$ as in (15). Replacing this value in the help IC gives

$$\pi^*=\frac{(1-\eta)2c+\frac{\eta\phi c}{3-\eta}}{\phi c+(1-\phi)B}. \tag{A.11}$$

When $\phi=0$, we have π^* as in (15). Notice that π^* is increasing in ϕ given Assumption 4 ($B\geq 2c$):

$$\frac{d\pi^*}{d\phi}=c\frac{\frac{\eta B}{3-\eta}-2(1-\eta)(c-B)}{[\phi c+(1-\phi)B]^2}>0.$$

If $\phi=1$, trivially the board does not need to pay for the CEO to exert effort, because the AI agent always finds a deal. Thus, the effort IC constraints are irrelevant in this case. By continuity, there exists $\overline{\phi}<1$ such that it is optimal to ignore all effort IC constraints for $\phi\geq\overline{\phi}$. If all second-period bonuses are zero, we have that the help IC constraint implies

$\pi^* = \frac{(1-\eta)w_1}{B}$. Then, we have that the derivative of V with respect to w_1 is

$$\frac{dV}{dw_1} = \phi\Big[\frac{M(1,\eta)}{B}\frac{1-\eta}{2} - \frac{2-\eta}{2}\Big],$$

which is negative because (13) holds. Thus, the optimal first period bonus is $w_1^* = 0$, implying that the optimal monitoring intensity is also $\pi^* = 0$ for all $\phi \geq \overline{\phi}$. □

References

ACEMOGLU, D., D. AUTOR, AND S. JOHNSON (2026): "Building Pro-Worker Artificial Intelligence," Tech. Rep. 34854, National Bureau of Economic Research.

ADAMS, R. B., A. C. AKYOL, AND P. VERWIJMEREN (2018): "Director Skill Sets," *Journal of Financial Economics*, 130, 641–662.

ADAMS, R. B. AND D. FERREIRA (2007): "A Theory of Friendly Boards," *Journal of Finance*, 62, 217–250.

AGHION, P. AND J. TIROLE (1997): "Formal and Real Authority in Organizations," *Journal of Political Economy*, 105, 1–29.

ALONSO, R. AND N. MATOUSCHEK (2008): "Optimal Delegation," *Review of Economic Studies*, 75, 259–293.

AMADOR, M. AND K. BAGWELL (2013): "The Theory of Optimal Delegation with an Application to Tariff Caps," *Econometrica*, 81, 1541–1599.

ATHEY, S. C., K. A. BRYAN, AND J. S. GANS (2020): "The Allocation of Decision Authority to Human and Artificial Intelligence," *AEA Papers and Proceedings*, 110, 80–84.

BABINA, T., A. FEDYK, A. HE, AND J. HODSON (2025): "Firm Investments in Artificial Intelligence Technologies and Changes in Workforce Composition," in *Technology, Productivity, and Economic Growth*, ed. by S. Basu, L. Eldridge, J. Haltiwanger, and E. Strassner, University of Chicago Press, vol. 83 of *Studies in Income and Wealth*.

BALDENIUS, T., N. MELUMAD, AND X. MENG (2014): "Board Composition and CEO Power," *Journal of Financial Economics*, 112, 53–68.

BALDENIUS, T., X. MENG, AND L. QIU (2019): "Biased Boards," *The Accounting Review*, 94, 1–27.

BARANCHUK, N. AND P. H. DYBVIG (2009): "Consensus in Diverse Corporate Boards," *The Review of Financial Studies*, 22, 715–747.

BURKART, M., D. GROMB, AND F. PANUNZI (1997): "Large Shareholders, Monitoring, and the Value of the Firm," *The Quarterly Journal of Economics*, 112, 693–728.

BURKART, M., S. MIGLIETTA, AND C. OSTERGAARD (2023): "Why Do Boards Exist? Governance Design in the Absence of Corporate Law," *The Review of Financial Studies*, 36, 1788–1836.

CHAKRABORTY, A. AND B. YILMAZ (2017): "Authority, Consensus, and Governance," *The Review of Financial Studies*, 30, 4267–4316.

CHEMMANUR, T. J. AND V. FEDASEYEU (2018): "A Theory of Corporate Boards and Forced CEO Turnover," *Management Science*, 64, 4798–4817.

CHEN, J. J. (2024): "Optimal Managerial Authority," Working paper, Boston University.

CHEN, J. J. AND B. Y. HAN (2026): "When Corporate AI Adoption Backfires," Working Paper, Boston University and University of Maryland.

CHEN, Y., J. GONG, J. LI, AND Z. ZHAO (2025): "Better Technology, Worse Motivation: GenAI's Mediocrity Trap," Available at SSRN: `https://ssrn.com/abstract=5208163`.

CHENG, X., M. DOGAN, AND P. YILDIRIM (2025): "Artificial Intelligence in Team Dynamics: Who Gets Replaced and Why?" Working paper.

COLES, J. L., N. D. DANIEL, AND L. NAVEEN (2008): "Boards: Does One Size Fit All?" *Journal of Financial Economics*, 87, 329–356.

COWGILL, B., P. HERNANDEZ-LAGOS, AND N. L. WRIGHT (2024): "Does AI cheapen talk? Theory and evidence from global entrepreneurship and hiring," *Theory and Ev-*

idence From Global Entrepreneurship and Hiring (July 26, 2024). Columbia Business School Research Paper.

CSASZAR, F. A., H. KETKAR, AND H. KIM (2024): "Artificial Intelligence and Strategic Decision-Making: Evidence from Entrepreneurs and Investors," *Strategy Science*, 9, 322–345.

DE HAAS, R., D. FERREIRA, AND T. KIRCHMAIER (2021): "The Inner Workings of the Board: Evidence from Emerging Markets," *Emerging Markets Review*, 48, 100777.

DELL'ACQUA, F., E. MCFOWLAND III, E. R. MOLLICK, H. LIFSHITZ-ASSAF, K. KELLOGG, S. RAJENDRAN, L. KRAYER, F. CANDELON, AND K. R. LAKHANI (2023): "Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality," Tech. Rep. Working Paper No. 24-013, Harvard Business School Technology & Operations Mgt. Unit, the Wharton School Research Paper. Available at SSRN: `https://ssrn.com/abstract=4573321`.

DEMIRER, M., J. J. HORTON, N. IMMORLICA, B. LUCIER, AND P. SHAHIDI (2026): "Chaining Tasks, Redefining Work: A Theory of AI Automation," Tech. Rep. 34859, National Bureau of Economic Research.

DESSEIN, W. (2002): "Authority and Communication in Organizations," *Review of Economic Studies*, 69, 811–838.

DONALDSON, J. R., N. MALENKO, AND G. PIACENTINO (2020): "Deadlock on the Board," *The Review of Financial Studies*, 33, 4445–4488.

DRYMIOTES, G. AND K. SIVARAMAKRISHNAN (2021): "Strategic Director Appointments," *Journal of Accounting Research*, 59, 1303–1347.

EWENS, M. AND N. MALENKO (2024): "Board Dynamics over the Startup Life Cycle," *Journal of Finance*, forthcoming.

FALEYE, O., R. HOITASH, AND U. HOITASH (2011): "The Costs of Intense Board Monitoring," *Journal of Financial Economics*, 101, 160–181.

FAMA, E. F. AND M. C. JENSEN (1983): "Separation of Ownership and Control," *The Journal of Law and Economics*, 26, 301–325.

FERREIRA, D., M. A. FERREIRA, AND B. MARIANO (2018): "Creditor Control Rights and Board Independence," *Journal of Finance*, 73, 2385–2422.

FERREIRA, D., M. A. FERREIRA, AND C. RAPOSO (2011): "Board Structure and Price Informativeness," *Journal of Financial Economics*, 99, 523–545.

FERREIRA, D. AND R. NIKOLOWA (2023): "Talent Discovery and Poaching Under Asymmetric Information," *Economic Journal*, 133, 201–234.

FIELD, L., M. LOWRY, AND A. MKRTCHYAN (2013): "Are Busy Boards Detrimental?" *Journal of Financial Economics*, 109, 63–82.

FUCHS, W., L. GARICANO, AND L. RAYO (2015): "Optimal Contracting and the Organization of Knowledge," *The Review of Economic Studies*, 82, 632–658.

GARICANO, L. (2000): "Hierarchies and the Organization of Knowledge in Production," *Journal of Political Economy*, 108, 874–904.

GARICANO, L., J. LI, AND Y. WU (2026): "Messy Jobs: The Work that AI Cannot Reach," Unpublished book manuscript.

GARICANO, L. AND E. ROSSI-HANSBERG (2006): "Organization and Inequality in a Knowledge Economy," *The Quarterly Journal of Economics*, 121, 1383–1435.

GONZALEZ-URIBE, J. AND M. GROEN-XU (2017): "CEO Contract Horizon and Innovation," Available at SSRN: `https://ssrn.com/abstract=2633763`.

GREENWALD, B. C. (1986): "Adverse Selection in the Labour Market," *Review of Economic Studies*, 53, 325–347.

HARFORD, J. AND R. J. SCHONLAU (2013): "Does the Director Labor Market Offer Ex Post Settling-up for CEOs? The Case of Acquisitions," *Journal of Financial Economics*, 110, 18–36.

HARRIS, M. AND A. RAVIV (2008): "A Theory of Board Control and Size," *The Review of Financial Studies*, 21, 1797–1832.

HOLMSTRÖM, B. (1984): "On the Theory of Delegation," in *Bayesian Models in Economic Theory*, ed. by M. Boyer and R. E. Kihlstrom, Amsterdam: North-Holland, 115–141.

HUANG, J. AND S. OUYANG (2025): "Soft Information, Hard Decisions: AI Advising," Working Paper.

IDE, E. AND E. TALAMÀS (2025): "Artificial Intelligence in the Knowledge Economy," *Journal of Political Economy*, 133, 3762–3800.

INDERST, R. AND H. M. MUELLER (2010): "CEO Replacement Under Private Information," *The Review of Financial Studies*, 23, 2935–2969.

JIANG, W. (2025): "Corporate Finance and Governance with Artificial Intelligence: Old and New," in *Artificial Intelligence in Finance*, ed. by T. Foucault, L. Gambacorta, W. Jiang, and X. Vives, CEPR Press, chap. 4.

JIANG, X. AND V. LAUX (2023): "What Role Do Boards Play in Companies with Visionary CEOs?" *Journal of Accounting Research*, 61, 1757–1796.

KENNY, G., M. KOWALKIEWICZ, AND K. OOSTHUIZEN (2025): "How CEOs Are Using Gen AI for Strategic Planning," *Harvard Business Review Digital Article*, online.

KORINEK, A. (2025): "AI Agents for Economic Research," NBER Working Paper 34202, National Bureau of Economic Research.

LARCKER, D. F., A. SERU, B. TAYAN, AND L. YOLER (2025): "The Artificially Intelligent Boardroom," Tech. Rep. 1, Stanford Graduate School of Business.

LEVIT, D. (2012): "Expertise, Structure, and Reputation of Corporate Boards," SSRN Working Paper.

——— (2020): "Words Speak Louder Without Actions," *Journal of Finance*, 75, 91–131.

LEVIT, D. AND N. MALENKO (2016): "The Labor Market for Directors and Externalities in Corporate Governance," *Journal of Finance*, 71, 775–808.

LI, J. (2013): "Job Mobility, Wage Dispersion, and Technological Change: An Asymmetric Information Perspective," *European Economic Review*, 60, 105–126.

LINCK, J. S., J. M. NETTER, AND T. YANG (2008): "The Determinants of Board Structure," *Journal of Financial Economics*, 87, 308–328.

MALENKO, N. (2014): "Communication and Decision-Making in Corporate Boards," *The Review of Financial Studies*, 27, 1486–1532.

——— (2024): "Information Flows, Organizational Structure, and Corporate Governance," in *Handbook of Corporate Finance*, ed. by D. Denis, Edward Elgar Publishing, chap. 14, 511–547.

MELUMAD, N. D. AND T. SHIBANO (1991): "Communication in Settings with No Transfers," *RAND Journal of Economics*, 22, 173–198.

NIKOLOWA, R. (2015): "Career Dynamics and Span of Control," *Economics Letters*, 128, 6–8.

RAHEJA, C. G. (2005): "Determinants of Board Size and Composition: A Theory of Corporate Boards," *Journal of Financial and Quantitative Analysis*, 40, 283–306.

RANSBOTHAM, S., D. KIRON, S. KHODABANDEH, S. IYER, AND A. DAS (2025): "The Emerging Agentic Enterprise: How Leaders Must Navigate a New Age of AI," *MIT Sloan Management Review and Boston Consulting Group*.

SHEKSHNIA, S. AND V. YAKUBOVICH (2025): "How Pioneering Boards Are Using AI," *Harvard Business Review*, 103, 56–64.

SONG, F. AND A. V. THAKOR (2006): "Information Control, Career Concerns, and Corporate Governance," *Journal of Finance*, 61, 1845–1896.

WALDMAN, M. (1984): "Job Assignments, Signalling, and Efficiency," *RAND Journal of Economics*, 15, 255–267.

WARTHER, V. A. (1998): "Board Effectiveness and Board Dissent: A Model of the Board's Relationship to Management and Shareholders," *Journal of Corporate Finance*, 4, 53–70.

YOTZOV, I., J. M. BARRERO, N. BLOOM, P. BUNN, S. J. DAVIS, K. M. FOSTER, A. JALCA, B. H. MEYER, P. MIZEN, M. A. NAVARRETE, P. SMIETANKA, G. THWAITES, AND B. Z. WANG (2026): "Firm Data on AI," Working Paper 34836, National Bureau of Economic Research, Cambridge, MA.

ZHONG, H. (2025): "Optimal Integration: Human, Machine, and Generative AI," *Management Science*, forthcoming.