

Online Appendix for

Board bias, information, and investment efficiency

Review of Accounting Studies

1 Project search

In this supplemental analysis, we extend our main results to a scenario where management is not restricted to the first encountered project. At date 1 (before the report is generated or the project value is observed), the CEO has a cost-free option to make additional independent draws from the pool of investment opportunities and search until she finds a project of the category (routine or novel) that she prefers.¹

The board's approval and the CEO's reporting and presentation decisions remain as before. At date 1, the CEO searches for a novel project if her margin from a novel project, $M(\theta_B) \equiv \mathbb{E}[(\theta - \theta_C)\mathbb{1}_{\theta \geq R(\theta_B)} | novel] - \mathbb{E}[(\theta - \theta_C)\mathbb{1}_{\theta \geq \max\{\theta_C, \theta_B\}} | routine]$, is at least zero and searches for a routine project otherwise.

Proposition OA-1 (Optimal board with first-order stochastic dominance)

(i) *If F and G are identical or if F first-order stochastically dominates G , then $\theta_B^* = \theta_H$ and the CEO pursues a novel project.*

(ii) *If G first-order stochastically dominates F :*

- *when $M(\theta_S) < 0$, then $\theta_B^* = \theta_S$ and the CEO pursues a routine project;*
- *when $M(\theta_S) \geq 0$, then $\theta_B^* \neq \theta_S$ and $\theta_B^* \notin [\theta_{min}, \theta_C]$.*

Because of her ability to opportunistically choose the reporting cutoff of novel but not routine projects, the CEO is more willing than the shareholders to pursue novel projects, all else

¹This is equivalent to a single draw with setting the probability p to be either zero (when novel project is preferred) or one (when routine project is preferred).

equal.² As a result, when the values of the projects follow the same distribution or when the value of novel projects (first-order) stochastically dominates that of routines, as is the case in Proposition OA-1 (i), the preferences of the board and the CEO about project category are aligned at $\theta_B = \theta_H$. Thus the optimal board has a bias of $\theta_B^* = \theta_H$, and the CEO pursues a novel project. Because the board can fully undo the CEO's empire-building, the equilibrium investment is efficient. When the value of routine projects stochastically dominates that of novel ones, two scenarios may arise. If the players' preferences at $\theta_B = \theta_S$ about project selection are aligned, as is the case of Proposition OA-1 (ii), the CEO searches for a routine project, the optimal board is neutral, $\theta_B^* = \theta_S$, and the equilibrium investment is efficient. However, if $M(\theta_S) \geq 0$ so that the preferences at $\theta_B = \theta_S$ are misaligned, the optimal board can never be neutral. This is intuitive: when faced with a neutral board, the CEO pursues a novel project. However, conditional on the project being novel, a neutral board is suboptimal for the shareholders.

So what is the optimal board in scenario (ii) with $M(\theta_S) \geq 0$? There are two options. First, the shareholders could acquiesce to the CEO's tendency to pursue a novel project and nominate a conservative board to counteract empire-building. Second, the shareholders could persuade the CEO to search for a routine project by nominating a board whose bias makes the pursuit of routine projects attractive from the CEO's perspective. The best option depends on the relative magnitude of the shareholders' benefit from pursuing a routine (stochastically dominating) project rather than a novel (stochastically dominated) one, and the shareholders' loss from investment distortions by a board that is biased enough to influence the CEO's project selection. Due to the generality of our model, there is little we can say about the board that sways the CEO into pursuing a routine project, apart from its not being neutral (as argued above) or having a stronger pro-expansion bias than the CEO (as this introduces investment distortions that can be mitigated by a different board).³

²The CEO's margin is $M(\theta_B) = W^{novel}(\theta_B) - W^{routine}(\theta_B) + (\theta_S - \theta_C)[G(\max\{\theta_B, \theta_C\}) - F(R(\theta_B))]$. The misalignment of preferences about project selection arises because of the third term, which reflects the difference in probability of obtaining a private benefit $\theta_S - \theta_C$ by the CEO.

³One could design conditions under which the optimal board has a mild pro-expansion bias to sway the CEO into selecting a routine project. This, for example, can occur if (i) $M(\theta) > 0$ for any $\theta \in [\theta_S, \theta_H]$; and (ii) $W^{routine}(\theta_C) > \max\{W^{novel}(\theta_H), W^{routine}(\theta_H)\}$ hold simultaneously. In this case, the optimal board has $\theta_B^* \in (\theta_C, \theta_S)$, and the CEO selects a routine project in equilibrium. To see why, note that the imposed condition (i) rules out the possibility that the CEO selects a routine project when faced with $\theta_B \in [\theta_S, \theta_H]$.

2 Alternative report structures

In the main model we focus on a binary reporting structure where $r = l$ is reported when $\theta \in [\theta_{min}, \theta_R]$ and $r = h$ is reported when $\theta \in [\theta_R, \theta_{max}]$. There, we briefly explain that this is one of the optimal reporting structures and no other structure can achieve a strictly better outcome for the CEO. We now show that this is indeed the case.

When the model parameters are such that the CEO chooses $\theta_R = \theta_C \geq H(\theta_B)$, the proof is trivial since the binary report implements the CEO's first-best outcome. Let us now focus on the case where the model parameters are such that the CEO sets $\theta_R = H(\theta_B) > \theta_C$. In this case the board is exactly indifferent over project acceptance and rejection when $r = h$.

Consider an alternative reporting structure where the continuum $[\theta_{min}, \theta_{max}]$ is split into n adjacent intervals indexed in an ascending order $i = 1, \dots, n$. A report $r = i$ is realized when $\theta \in [\underline{\theta}_i, \bar{\theta}_i]$. The CEO chooses the number of intervals n and the corresponding cutoffs $\underline{\theta}_i$ and $\bar{\theta}_i$. The posterior project value conditional on report $r = i$ is increasing in i , and therefore the reports can be divided into reports that induce $a = 0$ (the project is rejected because the posterior is insufficient, $\mathbb{E}[\theta | \theta \in [\underline{\theta}_i, \bar{\theta}_i]] < \theta_B$) and reports that induce $a = 1$ (the project is accepted because the posterior is sufficient, $\mathbb{E}[\theta | \theta \in [\underline{\theta}_i, \bar{\theta}_i]] \geq \theta_B$).

We prove that this information structure cannot be better for the CEO. First of all, see that the posterior project value on the interval $[\underline{\theta}_i, \bar{\theta}_i]$, $\mathbb{E}[\theta | \theta \in [\underline{\theta}_i, \bar{\theta}_i]]$, is increasing in both $\underline{\theta}_i$ and $\bar{\theta}_i$. Take the *lowest* report out of all reports that generate a sufficiently large posterior and denote it k . (If no such report exists, the information structure is clearly worse, as the project is rejected for any θ .) If the richer information structure is better for the CEO than the binary structure, then we must have $\underline{\theta}_k < \theta_R = H(\theta_B)$. Since $\underline{\theta}_k < \theta_R$ and $\bar{\theta}_k \leq \theta_{max}$, the board's posterior project value for a report k in the richer structure is *lower* than the board's project value for a report h in the binary structure, $\mathbb{E}[\theta | \theta \in [\underline{\theta}_k, \bar{\theta}_k]] < \mathbb{E}[\theta | \theta \in [\theta_R, \theta_{max}]] = \theta_B$. Since $\mathbb{E}[\theta | \theta \in [\underline{\theta}_k, \bar{\theta}_k]] < \theta_B$, the board that observes $r = k$ rejects the project. This is a contradiction.

Finally, while a richer information structure cannot give the CEO a higher payoff, it can

Moreover, the imposed condition (ii) rules out the optimality of any board with $\theta_B > \theta_H$ but also ensures that the board prefers to sway the CEO into selecting a routine project. Lastly, by Proposition [OA-1](#), the board cannot have a (weakly) stronger pro-expansion bias than that of the CEO.

give an identical payoff. First, it is always possible to refine the low report $r = l$ by setting any number of intervals for the continuum $[\theta_{min}, \theta_R^*]$, but the CEO will still decide to not present the project with the refined (more precise) information. Second, when $\theta_R = \theta_C > H(\theta_B)$, it is possible to refine $r = h$ as long as the lowest report out of all reports that generate a sufficiently large posterior satisfies $\underline{\theta}_k = \theta_C$. When $\theta_R = H(\theta_B) \geq \theta_C$, it is never optimal to refine the high report $r = h$. See also Friedman, Hughes, and Michaeli (2020) for weak optimality of binary report structures with threshold payoffs (similar to the ones obtained when the receiver takes a binary action as in our setting).

3 Cheap talk

In the main analysis, we solely focus on the CEO's collection of publicly observable information. In practice, after preparing a formal report for a board meeting, CEOs may come across additional information and informally communicate it to the board. In this extension, we briefly demonstrate that our main results are robust to accounting for such a possibility. In particular, we consider a setting where, after the public report is observed and the project is proposed, the CEO privately learns the novel project value and sends a non-verifiable message to the board (cheap talk communication). The timing of the public report realization does not matter.

Proposition OA-2 (Robustness of the results with cheap talk) *Suppose that, after the public report is observed and the project is proposed, the CEO privately learns the novel project value and sends a non-verifiable message to the board. Then, the CEO's equilibrium non-verifiable message is uninformative, and the equilibrium reporting cutoff is $\theta_R^* = R(\theta_B)$ as defined in Lemma 4 of the main text.*

In equilibrium, the post-report cheap talk communication is uninformative and has no effect on the CEO's reporting strategy.⁴ Put differently, the presence of cheap talk option preserves the communicated hard information without conveying additional soft information. Our analysis implies that the CEO at least weakly prefers committing to verifiable messages

⁴Essentially, the CEO's message only confirms that the project value belongs to the subinterval associated with the realized report.

and opting out of subsequent non-verifiable communication. Given Proposition OA-2, the findings in the main model remain intact even when the CEO privately learns the novel project value (after setting the reporting cutoff) and sends a non-verifiable message to the board.⁵

4 Private reports and misreporting

We now consider an extension where the CEO privately observes the report r about a novel project and communicates a message $m_r \in \{l, r\}$ to the board. If the CEO misreports, i.e., $m_r \neq r$, she incurs a non-negative cost $\gamma \geq 0$.

We note that if $\theta_B < \theta_C$, the CEO obtains her most preferred outcome by setting $\theta_R = \theta_C$ and presenting only projects with $r = h$. The possibility to misreport in this case has no impact, as the CEO, having observed any r , has no incentives to deviate from $m_r = r$ and $(d_l, d_h) = (0, 1)$.

Next we turn to the more interesting case, $\theta_B \geq \theta_C$. If $\theta_B \leq \mathbb{E}[\theta]$, the board that believes $r = h$ approves the project for any reporting cutoff $\theta_R \geq \theta_C$.⁶ However, if $\theta_B \geq \mathbb{E}[\theta]$, ensuring the board's approval when the board believes $r = h$ imposes a lower bound $\theta_R \geq H(\theta_B)$ on the reporting cutoff. Lemma 4 in the main text showed that, with public reports, this is the only constraint. When the report is private, the board still has to be motivated to approve a presented project, $\theta_R \geq H(\theta_B)$, but the CEO's communication of $m_r = r$ is no longer guaranteed. There are two candidates for an equilibrium in the subgame starting at date 2. The first is a separating equilibrium where a CEO who has observed $r = h$ sends $m_h = h$ and presents the project, and a CEO who has observed $r = l$ sends $m_l = l$ and does not present the project. If this separating equilibrium fails to exist due to deviation of any type, then the only equilibrium in the subgame starting at date 2 is a pooling equilibrium. In this pooling equilibrium, the board ignores all CEO messages and approves or rejects the

⁵If we additionally alter the main model such that the CEO *cannot choose the properties of the publicly observable report* and can only send a non-verifiable message to a board that takes a binary action (accept/reject), the optimal board from shareholders' perspective will be neutral. Analysis available upon request.

⁶Lower reporting cutoffs are irrelevant as the conflict of interests between the CEO and the shareholders exists only on the interval $\theta_R \in [\theta_C, \theta_S]$.

proposed project based on its prior beliefs, $\mathbb{E}[\theta] \leq \theta_B$.

When $\theta_R \geq \theta_C$, a CEO who has observed $r = h$ strictly prefers to communicate $m_h = h$. But a CEO who has observed $r = l$ may misreport. There exists a reporting cutoff at which, following $r = l$, the CEO is indifferent between project rejection (via costless message $m_l = l$) and project approval (via costly misreporting $m_l = h$). This indifference writes

$$\mathbb{E}[\theta | \theta \leq \theta_R] - \theta_C - \gamma = 0$$

and is equivalent to the board's indifference over the two outcomes, where the board type is $\theta_B = \theta_C + \gamma$. Using our established notation, this indifference holds for a reporting cutoff $\theta_R = L(\theta_C + \gamma)$. If the reporting cutoff is higher, the expected project value after $r = l$ is too attractive for a CEO who has observed $r = l$, and she misreports. Therefore, the CEO of type $r = l$ is willing to communicate $m_l = l$ and withhold the project, $d_l = 0$, only if the reporting cutoff complies with an upper bound

$$\theta_R \leq L(\theta_C + \gamma).$$

Note that $\theta_C < L(\theta_C) \leq L(\theta_C + \gamma)$. Therefore, only two situations exist:

- (i) If $\max\{\theta_C, H(\theta_B)\} \leq L(\theta_C + \gamma)$, the upper bound is loose at $\theta_R^* = \max\{\theta_C, H(\theta_B)\}$. An informative equilibrium with a reporting cutoff $\theta_R^* = \max\{\theta_C, H(\theta_B)\}$, messages $(m_l, m_h) = (l, h)$, presentation decisions $(d_l, d_h) = (0, 1)$, and approval $a_h = 1$ exists.
- (ii) If $\theta_C < L(\theta_C + \gamma) < H(\theta_B)$, there is no reporting cutoff θ_R that simultaneously satisfies the lower and the upper bound. The equilibrium is uninformative. Because this situation occurs only when the lower bound restricts the CEO's implementation of cutoff θ_C (i.e., when $H(\theta_B) > \theta_C$), we must have that $H(\theta_B) > \theta_C > \theta_{min}$, which implies $\theta_B \geq \mathbb{E}[\theta]$ (see also Figure 3), and so after an uninformative message the board *rejects* the project.

We next consider the board type that maximizes firm value *if the project is novel with certainty*. In the main setting, this board has type $\theta_B^* = \theta_H$. We are interested in finding how the introduction of private reports changes this prediction. We assume that whenever

the shareholders are indifferent between multiple board types, they choose the highest one. As our result below shows, the key consideration is whether the upper bound on θ_R is binding for a board of type $\theta_B = \theta_H$ or not. For the sets of parameters where this constraint is loose, the introduction of private reports does not change the outcome, $\theta_B^* = \theta_H$, and investments are efficient. For the set of parameters where this constraint is binding, the optimal board is still conservative but can be either more mildly biased $\theta_B^* \in (\theta_S, \theta_H)$ or extremely biased $\theta_B^* = \theta_{max} > \theta_H$, as per our tie-breaking rule. Consequently, we either observe overinvestment or extreme underinvestment (all projects are rejected).

Proposition OA-3 *Suppose that the project is novel with certainty. The optimal board from a shareholders' perspective is conservative, $\theta_B^* \in (\theta_S, \theta_{max}]$. In particular:*

- *If $L(\theta_C + \gamma) \geq \theta_S$, then $\theta_B^* = \theta_H$. The investment in the novel project is efficient.*
- *If $L(\theta_C + \gamma) < \theta_S$, then:*
 - (i) *$\theta_B^* = H^{-1}(L(\theta_C + \gamma)) \in (\theta_S, \theta_H)$ if $\theta_S \leq \mathbb{E}[\theta]$ or $\theta_S > \mathbb{E}[\theta]$ and $L(\theta_C + \gamma) \geq H(\theta_S)$. The investment in the novel project is mildly excessive.*
 - (ii) *$\theta_B^* = \theta_{max} > \theta_H$ if $\theta_S > \mathbb{E}[\theta]$ and $L(\theta_C + \gamma) < H(\theta_S)$. The novel project is rejected.*

The preceding discussion focused on the case where the project is novel with certainty. Given that the optimal board then is still conservative, accounting for the possibility that the project is routine with some probability will not qualitatively change our results from the main setting.

5 Board's information acquisition with unknown cost

In this supplemental analysis, we extend Section 5.1 of the main text by studying the case where κ is unknown to the CEO and is revealed to the board when it chooses θ_M on date 3. We assume that κ is drawn from some distribution with support $[\underline{\kappa}, \bar{\kappa}]$, where $\underline{\kappa} > 0$. Given that the board observes the cost when acquiring information, its optimal message cutoff is the same as the one in Lemma 4 in the main text.

Because the CEO does not observe κ , she anticipates a range of possible realizations. It turns out that her optimization across these realizations is simple: any message cutoff that is optimal for $\kappa = \bar{\kappa}$ is also optimal for $\kappa < \bar{\kappa}$, i.e., for all possible cost realizations and distributions.

Lemma OA-1 *At date 2, the CEO sets $\theta_R \in [R(\theta_B), \hat{R}(\theta_B, \bar{\kappa})]$, where $R(\theta_B) = \max\{\theta_C, H(\theta_B)\}$ as defined in Lemma 4 of the main text and $\hat{R}(\theta_B, \bar{\kappa}) \equiv \max\{\theta_C, H(\theta_B), \theta_B - \bar{\kappa}\}$. At date 3, the CEO presents the novel project with $r = h$, the board observes κ and sets $\theta_M = \hat{R}(\theta_B, \kappa) \geq \theta_R$. At date 4, the board approves projects with $\theta \geq \theta_M$ and rejects the rest.*

From the shareholders' point of view, the message cutoff spreads across a range, $\theta_M \in [\hat{R}(\theta_B, \bar{\kappa}), \hat{R}(\theta_B, \underline{\kappa})]$.⁷ As for the board type that invests efficiently in novel investments, the following tradeoff arises: a strongly conservative board ensures efficient investment in novel projects when κ is high but underinvests when it is low. And vice versa, a mildly conservative board invests efficiently when κ is small but overinvests when it is high. Balancing these distortions in the optimum, the board with $\hat{\theta}_H$ overinvests when the cost is high and underinvests when the cost is low.

Lemma OA-2 *It holds that $\theta_S < \min\{\theta_H, \theta_S + \underline{\kappa}\} \leq \hat{\theta}_H \leq \min\{\theta_H, \theta_S + \bar{\kappa}\}$.*

So what is the optimal board type from shareholders' perspective that takes into account that projects can be routine or novel? Adding uncertainty over information costs has no effect on the existence of a tradeoff between a neutral and a conservative board. A neutral board invests efficiently in routines but overinvests in novel projects. A conservative board underinvests in routines; for novel projects, it overinvests if the realized board cost is sufficiently large. When κ is small, it may happen that the board underinvests in novel projects.

⁷If even the lowest cost realization is prohibitively high, the interval for reporting cutoff degenerates to a single value $R(\theta_B)$, and the board that invests efficiently in novel projects is θ_H , as in the main analysis.

6 Boards' information acquisition with entropy cost

This supplemental analysis illustrates that the main result about the benefit from conservative boards remains robust to assuming entropy-based information acquisition cost and full board's signal space. To streamline the discussion, we fix $d = 1$, focus solely on novel projects, and assume a trinary project value, $\theta \in \Theta \equiv \{\theta_1, \theta_2, \theta_3\}$ with $\theta_1 < \theta_2 < 0 < \theta_3$. As before, $\theta_S = 0$; i.e., the firm value is strictly positive only if the project yielding the highest value, θ_3 , is approved (i.e., only if $\theta = \theta_3$ and $a = 1$). We assume that $\theta_C \in (\theta_1, \theta_2)$, which implies that the CEO receives a positive payoff not only if the project of the highest value, θ_3 , is approved but also if the one of intermediate value, θ_2 , is approved (i.e., when $\theta \in \{\theta_2, \theta_3\}$ and $a = 1$). Put differently, the CEO is biased in favor of adopting the intermediate project (her empire-building preference). For the board, it is sufficient to focus on $\theta_B \in (\theta_2, \theta_3)$.⁸

The prior probability distribution of θ is common knowledge. Formally, the prior belief is given by $\mu^o \equiv (\mu_1^o, \mu_2^o, \mu_3^o) \in \Delta(\Theta)$, where $\mu_j^o \equiv \Pr(\theta = \theta_j)$. The CEO chooses the properties of a verifiable public report R about the project value from a finite set $\mathbb{S}$ of possible realizations. For our analysis, it is useful to denote $\phi^r \equiv \Pr(R = r)$ the probability that the realized report is r . A report realization $r \in \mathbb{S}$ induces an interim belief $\mu^r \equiv (\mu_1^r, \mu_2^r, \mu_3^r)$, where $\mu_j^r \equiv \Pr(\theta = \theta_j \mid r)$, $j \in \{1, 2, 3\}$. To avoid clutter, we occasionally suppress the superscript r and simply use μ . Only when the realized report r becomes relevant for the analysis do we use μ^r . Because every report realization r is associated with a specific belief μ^r , the distribution $\phi = (\phi^r)_{r \in \mathbb{S}}$ also describes the (probability) distribution of the interim beliefs.⁹ The CEO can select any ϕ with a finite number of report realizations as long as it satisfies the martingale (Bayes-plausibility) property, $\mathbb{E}_\phi[\mu] = \mu^o$. The distribution ϕ and the posterior beliefs μ jointly determine the precision of the report.¹⁰

After observing the realization r and forming interim belief μ^r , the board may obtain an additional signal T from a set $\mathbb{T}$ of a finite number of possible realizations. We denote

⁸No other board can be optimal for the shareholders: (i) A board with $\theta_B \leq \theta_1$ ($\theta_B \geq \theta_3$) approves (rejects) all projects for any information. (ii) A board with $\theta_B \in (\theta_1, \theta_2]$ is fully aligned with the CEO, receives a perfect report, and approves the intermediate (value-destroying) project.

⁹Specifically, ϕ^r is also the probability that the interim belief is μ^r , so there is a one-to-one mapping between the distribution of the report and the distribution of the beliefs that the report induces.

¹⁰In Lemma OA-3, we show that an optimally set binary report achieves the same expected utility for the CEO. But, for now, we allow more than two report realizations.

$\tau^t \equiv \Pr(T = t)$ the probability that the additional signal realization is t . The properties of the additional signal are characterized by $\tau = (\tau^t)_{t \in \mathbb{T}}$. For every realization $t \in \mathbb{T}$, there are corresponding final beliefs $\mu^t \equiv (\mu_1^t, \mu_2^t, \mu_3^t)$, where $\mu_j^t \equiv \Pr(\theta = \theta_j \mid r, t)$.¹¹ The martingale property, $\mathbb{E}_\tau[\mu^t] = \mu$, must hold.

Learning additional information is costly to the board. We assume that the board's cost of obtaining the additional signal is proportional to the reduction of the (expected) Shannon entropy so that learning a more informative signal is costlier for the board. Formally, the Shannon entropy of the $|\Theta|$ -dimensional interim belief μ (for given r and ϕ before observing t) is given by $H(\mu) = -\sum_j \mu_j \ln \mu_j$, where $0 \ln 0 = 0$ holds by convention. The total entropy-based cost of a signal distributed by τ over a support $\mathbb{T}$ is $\sum_{t \in \mathbb{T}} \tau^t H(\mu^t)$. Then, the board's personal (entropy-based) cost of the signal T is increasing in the reduction of the expected entropy, $J(\mu, \tau) = \kappa \{H(\mu) - \sum_{t \in \mathbb{T}} \tau^t H(\mu^t)\}$, where $\kappa \geq 0$ is the marginal cost of reducing the entropy.¹² The lower κ , the easier it is for the board to learn information.

Board learning and project decision. We solve this model variation by backward induction. At date 5, for given report $r \in \mathbb{S}$ and additional signal $t \in \mathbb{T}$, the board approves the project if approval yields a higher expected payoff than rejection:

$$\mathbb{E}[u_B(1, \theta) \mid r, t] \geq \mathbb{E}[u_B(0, \theta) \mid r, t] = 0. \quad (\text{OA-1})$$

The approval decision depends on the report and the signal. The information from the report and the signal is summarized in the final belief μ^t , so we let $a(\mu^t)$ denote the board's approval decision for the final belief μ^t . Formally, $a(\mu^t) = \mathbb{1}\{\sum_j \mu_j^t \theta_j - \theta_B \geq 0\}$, where $\mathbb{1}$ is an indicator function.

At date 4, after observing report $r \in \mathbb{S}$ and anticipating its approval strategy, the board decides whether and how much additional information to learn. Technically speaking, the

¹¹The finite belief μ^t is again a function of the realized report r . We suppress it to avoid clutter.

¹²To fix ideas, consider an example with intermediate beliefs $\mu^r = (1/3, 1/3, 1/3)$ so that the Shannon entropy is $H(\mu^r) = -(1/3 \ln 1/3 + 1/3 \ln 1/3 + 1/3 \ln 1/3) = \ln 3$. A perfectly informative additional signal eliminates all uncertainty and thereby results in the Shannon entropy of zero. Thus, the board's cost of acquiring such signal is $J(\mu, \tau) = \kappa(\ln 3 - 0) = \kappa \ln 3$. Alternatively, a completely uninformative private signal does not reduce any uncertainty, so the reduction in (expected) Shannon entropy is zero and the cost incurred by the board is $J(\mu, \tau) = \kappa(\ln 3 - \ln 3) = 0$. Any imperfectly informative additional signal will be associated with a cost $J(\mu, \tau) \in (0, \kappa \ln 3)$.

board's problem is to find a distribution τ that maximizes the board's conditionally expected payoff net of personal learning (entropy-based) costs,

$$\mathbb{E} [u_B(a(\mu^t), \theta) \mid r, \tau] - J(\mu, \tau),$$

subject to the martingale property. For every report realization r and corresponding intermediate belief μ^r , there is an optimal corresponding signal distribution $\tau_B(\mu^r)$ of the additional signal learned by the board.

As we show below, of specific interest for our analysis are the report realizations (and the associated beliefs) for which the board chooses *not to learn* any additional information (or, to put it differently, learns an uninformative signal). To study these report realizations we use a graphical argument. Because the state can assume three possible values, the board's beliefs are elements of a two-dimensional simplex $\Delta(\Theta)$. We denote its corners as $\{A, B, C\}$ with

$$\mu^A = (1, 0, 0); \quad \mu^B = (0, 1, 0); \quad \mu^C = (0, 0, 1).$$

Upon observing $r = A$, the board is certain that the project value is θ_1 . Similarly, after $r = B$ ($r = C$), the board believes that the project value is θ_2 (θ_3). For any possible report, the board's interim belief is within the two dimensional simplex and thus is a convex combination of the corner beliefs. We label the simplex $\Delta^{ABC} \equiv \Delta(\Theta)$. In addition, let $NL^a \subset \Delta^{ABC}$ represent the set of interim beliefs for which the board chooses not to learn any additional information (and thus incurs zero learning entropy-based costs) and takes an action a . We refer to NL^a as the “nonlearning region of action a .”

Lemma OA-3 *There exist two non-empty nonlearning regions, $NL^1 \subset \Delta^{ABC}$ and $NL^0 \subset \Delta^{ABC}$, such that:*

- (i) *If $\mu^r \in NL^1$, the board does not learn additional information and approves the project. For any properties of the report, the nonlearning region NL^1 of the board shrinks in the board's bias θ_B and expands in κ .*
- (ii) *If $\mu^r \in NL^0$, the board does not learn additional information and rejects the project. For any properties of the report, the nonlearning region NL^0 expands in the board's*

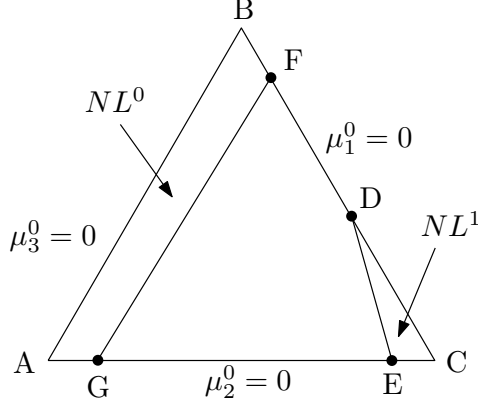


Figure 1: Interim beliefs and nonlearning regions in the simplex $\Delta(\Theta)$

bias θ_B and κ .

The non-learning region of approval, NL^1 , and the nonlearning region of project rejection, NL^0 , are graphically illustrated in Figure 1. Beliefs in the corners of the nonlearning regions, D , E , F , and G , are derived in Proof to Lemma OA-3.

With a higher board's bias θ_B , what matters for the size of the nonlearning regions are the ratios of the board's gain from the acceptable project, $\theta_3 - \theta_B$, to the board's losses from unacceptable projects, $\theta_B - \theta_2$ and $\theta_B - \theta_1$. With a higher board's bias θ_B , the ratios decrease, which implies that NL^1 shrinks and NL^0 expands. The comparative statics with respect to κ is straightforward: The costlier it is to learn additional information, the more often the board will choose to forgo learning (i.e., both regions expand).

CEO's optimal report. At date 2, the CEO chooses the properties of the report. Her problem is to find a distribution ϕ that maximizes her expected payoff subject to the martingale property. Kamenica and Gentzkow (2011) show that the optimal distribution can be obtained by concavification of the CEO's expected payoff function conditional on the observed report, i.e., the CEO's payoff function over the interim beliefs invoked by the report. Characterizing the CEO's payoff function in our setting, in the presence of the board's learning option, is nontrivial. Nevertheless, the solution of the problem can be presented elegantly. Similar to recent literature on generalized Bayesian persuasion of a rationally inattentive receiver (Matysková and Montes 2023; Caplin, Dean, and Leahy 2019), we find

that the CEO optimally (i) ensures the interim beliefs are such that the board does not learn additional information and (ii) chooses the distribution of interim beliefs that maximizes the probability of the board’s approval of intermediate and high-value projects. To understand why the CEO optimally avoids learning by the board, note that, while at date 2 the CEO is uncertain which report will be realized, she controls the set of possible supported realizations and the probability with which each one of these realizations is generated. Because the CEO is unconstrained in her choice and because the board’s information cost is additively separable in posteriors, any information that the board decides to acquire can also be directly provided by the CEO such that the board doesn’t acquire any other information. Then, the CEO controls the set of final posterior beliefs used as a basis for the board’s approval decision, as well as the ex ante distribution of these beliefs.¹³

Because the CEO optimally wants to ensure the board has no incentives to learn additional information, the candidate report realizations are in the nonlearning regions NL^0 and NL^1 . Our problem can be further simplified by focusing only on the extreme points of these nonlearning regions (Matysková and Montes 2023). An extreme point of NL^a is a point with belief $\mu \in NL^a$, which does not lie on any open line segment joining two points of NL^a . Focusing on the extreme points means considering only report realizations that imply the project is one of two types, i.e., realizations that rule out one state with certainty.¹⁴

In our case, the set of extreme points of NL^0 is $\{A, B, F, G\}$, and the set of extreme points of NL^1 is $\{C, D, E\}$. Thus, the set of candidate realizations that could be supported by the optimal report is

$$\widehat{\mathbb{S}} = \{A, B, C, D, E, F, G\}.$$

The properties of the report on the reduced set $\widehat{\mathbb{S}}$ are represented by a distribution

$$\phi = \{\phi^A, \phi^B, \phi^C, \phi^D, \phi^E, \phi^F, \phi^G\},$$

¹³Caplin, Dean, and Leahy (2019) demonstrate that the board’s optimal signal mixes final beliefs that are borders of non-learning regions. The CEO can invoke them directly by the choice of the reporting technology. In addition, however, the CEO may invoke other-than-border points from the nonlearning regions, and exactly this difference in our setting makes it strictly optimal for the CEO to discourage board learning.

¹⁴Technically, because the CEO’s expected payoff function is linear within non-learning regions, the concave closure cannot involve only points in the interior, and any value in the interior (implying that the project can be of any type) can be achieved through a lottery on the extreme points.

where, as before, $\phi^r \equiv \Pr(R = r)$, $r \in \widehat{\mathbb{S}}$. The optimal report is characterized by the optimal distribution $\tilde{\phi}$ over the reduced set $\widehat{\mathbb{S}}$, where $\tilde{\phi} = \left(\tilde{\phi}^r\right)_{r \in \widehat{\mathbb{S}}} = \arg \max_{\phi \in \Delta(\widehat{\mathbb{S}})} \mathbb{E}[u_C(\cdot)]$, subject to the martingale property, $\mu^o = \sum_{r \in \widehat{\mathbb{S}}} \phi^r \mu^r$. Lemma OA-4 below reduces the dimensionality of the problem by establishing that the optimal report properties do not support realizations $\{E, F, G\}$.

Lemma OA-4 *The optimal report is not supported by interim beliefs that lie on the points $\{E, F, G\}$, i.e., $\tilde{\phi}^E = \tilde{\phi}^F = \tilde{\phi}^G = 0$.*

The result in Lemma OA-4 implies that the optimal report can only be supported by realizations $r \in \{A, B, C, D\}$. Clearly, upon observing $r \in \{A, B, C\}$, the board is fully informed and commits no approval errors. As a consequence, the only realization supported by the optimal report that leaves the board not fully informed is $r = D$. This report implies that the project either yields high or intermediate value. For our subsequent analysis, it is instructive to describe the point D through its likelihood ratio

$$\omega(\theta_B) \equiv \frac{\Pr(\theta = \theta_3 \mid r = D)}{\Pr(\theta = \theta_2 \mid r = D)} = \frac{\mu_3^D}{1 - \mu_3^D}.$$

The larger the $\omega(\theta_B)$, the larger the probability that the project has a high value, conditional on observing $r = D$. Because, by construction, the board approves the project at point D and the only approval error occurs at point D , we label $\omega(\theta_B)$ the “approval precision” of the board. For future reference, we additionally label the likelihood ratio at the prior beliefs, $\omega^o \equiv \mu_3^o / \mu_2^o$, the “guaranteed” level of approval precision. It is also useful to introduce $\theta_B^o \in (\theta_2, \theta_3)$ as a board type satisfying $\omega^o = \omega(\theta_B^o)$, i.e., $\theta_B^o = \omega^{-1}(\omega^o)$.¹⁵

Corollary OA-1 *The board’s approval precision $\omega(\theta_B)$ is decreasing in the cost of learning, κ , and increasing in the board bias, θ_B .*

As a next step, we classify all possible parametrical cases into two types depending on the intensity of the conflict between the CEO and the board. Let Δ_{ABD} be the convex hull of $\{A, B, D\}$ and Δ_{ACD} be the convex hull of $\{A, C, D\}$. When the prior belief μ^o is in

¹⁵To see that $\theta_B^o > \theta_2$, we use that $\omega(\theta_2) = 0$ and $\omega^o > 0$. To see that $\theta_B^o < \theta_3$, we use that $\omega(\theta_3) \rightarrow \infty > \omega^o$. Notably, because μ_3^D does not depend on the properties ϕ , the cutoff θ_B^o is also independent of ϕ .

Δ_{ABD} , the project is more likely to have intermediate value. Because this project is value-destroying for the firm and its shareholders, we label this case “strong conflict.” It is easy to see that strong conflict is equivalent to a situation in which the approval precision exceeds the guaranteed level of precision, $\omega(\theta_B) > \omega^o$. Because the approval precision $\omega(\theta_B)$ is continuously increasing in θ_B and the guaranteed approval precision ω^o is independent of θ_B , we can also present the strong conflict as a situation in which $\theta_B > \theta_B^o$. Similarly, when the prior belief μ^o is in Δ_{ACD} , the project is more likely to have high value. Because this project adds value to the firm, we label this case “weak conflict” of interests.¹⁶ Following a similar argument, weak conflict is characterized by $\omega(\theta_B) < \omega^o$ or, equivalently, $\theta_B < \theta_B^o$.

We are now ready to solve for the optimal properties $\tilde{\Phi}$. Proposition OA-4 shows that the system depends on the severity of the conflict defined above, and is not fully informative as long as the projects could be of intermediate value, $\mu_2^o > 0$. The optimal CEO’s report is illustrated in Figure 2 (a).

Proposition OA-4 (Optimal properties of the CEO’s report)

- (i) For a strong conflict between the board and the CEO ($\theta_B > \theta_B^o$), the report is characterized by the optimal distribution $\tilde{\Phi} = \Phi_{ABD} \equiv \left(\mu_1^o, \mu_2^o - \frac{\mu_3^o(1-\mu_3^D)}{\mu_3^D}, 0, \frac{\mu_3^o}{\mu_3^D}, 0, 0, 0 \right)$. The board approves the project after observing $r = D$ and rejects the project otherwise.
- (ii) For a weak conflict between the board and the CEO ($\theta_B < \theta_B^o$), the report is characterized by the optimal distribution $\tilde{\Phi} = \Phi_{ACD} \equiv \left(\mu_1^o, 0, \mu_3^o - \frac{\mu_2^o\mu_3^D}{1-\mu_3^D}, \frac{\mu_2^o}{1-\mu_3^D}, 0, 0, 0 \right)$. The board rejects the project after observing $r = A$ and approves the project otherwise.

When the conflict between the board and the CEO is strong ($\theta_B > \theta_B^o$), only realizations $r \in \{A, B, D\}$ are supported. The project with the lowest (highest) value is always reported as $r = A$ ($r = D$). However, the intermediate type project is sometimes reported as $r = B$ and sometimes as $r = D$. Because, by construction, only $r = D$ is associated with project approval, the high-value and (with some probability) the intermediate-value projects are approved. When the conflict is weak ($\theta_B < \theta_B^o$), the optimal report only supports $r \in \{A, C, D\}$. The project with the lowest (intermediate) value is always reported as $r = A$

¹⁶Note that $\Delta_{ABD} \cup \Delta_{ACD} = \Delta(\Theta)$.

($r = D$). The high-value project is sometimes reported as $r = C$ and sometimes as $r = D$. Now the board approves the project following two reports, $r \in \{C, D\}$, and the high and intermediate-value projects are approved. The reason is that when the conflict is weak, the prior beliefs about the project value are high, so persuading the board to approve all intermediate-value projects by mixing them with some high-value projects is feasible.

In summary, a mild conflict of interest between the CEO and the board is associated with low θ_B . In this case, the board approves the intermediate- and high-value projects. Once the level of board bias reaches the critical cutoff θ_B^o , the board begins to reject some of the intermediate-value projects. Any increase in θ_B beyond that point further decreases the chance that the intermediate-value project is approved.

Corollary OA-2 (Probability of approval) *When the conflict is weak ($\theta_B < \theta_B^o$), the ex ante probability of project approval is $\Pr(a = 1) = \Pr(r = D) + \Pr(r = C) = \mu_2^o + \mu_3^o$. The shareholders' ex ante expected payoff is $\mu_3^o (\theta_3 + \theta_2/\omega^o)$. When the conflict is strong ($\theta_B > \theta_B^o$), the ex ante probability of project approval is given by $\Pr(a = 1) = \Pr(r = D) = \mu_3^o/\mu_3^D \in (\mu_3^o, \mu_2^o + \mu_3^o)$ and is decreasing in θ_B . The shareholders' ex ante expected payoff is given by $\mu_3^o (\theta_3 + \theta_2/\omega(\theta_B))$ and is increasing in θ_B .*

Benefit of board's conservatism. Regardless of the severity of the conflict, the optimal report properties do not distort the approval decision for the projects about which the CEO, the board, and the shareholders agree. The lowest-value projects are always rejected, and the highest-value projects are always approved. The only distortion to the company-value-maximizing decision is that some (or all) intermediate projects are also approved. This distortion tends to vanish when the board becomes more conservative and its conflict of interest with the CEO becomes more severe (by Lemma OA-3, a higher θ_B shifts the point D closer the point C), as the board is less likely to approve the intermediate-value project (see also the comparative statics in Corollary OA-1). It is therefore in the interest of the shareholders to *increase the board bias* such that the conflict between the board and the CEO is as strong as possible, but to also preserve $\theta_B < \theta_3$, to keep the board's preference on the highest-value project non-distorted. Formally, by Corollary OA-2, the shareholders' ex ante

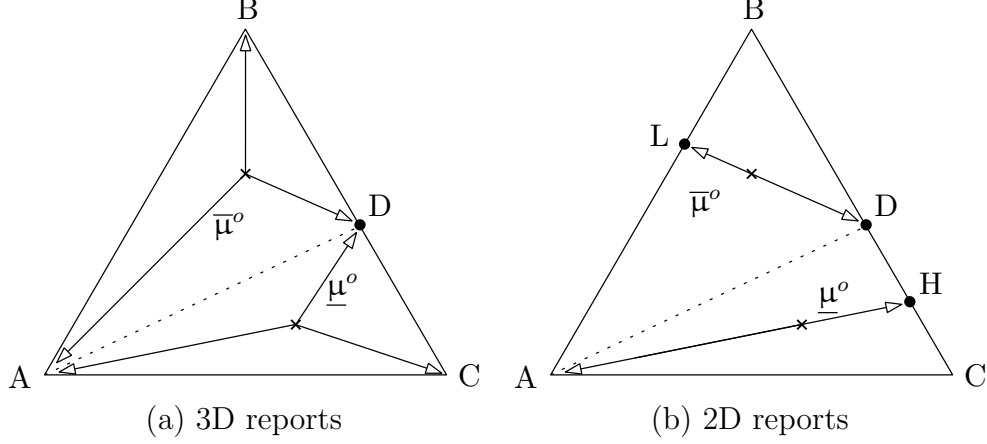


Figure 2: Equivalent characterizations of the optimal CEO's report
A weak conflict is represented by $\underline{\mu}^o$, and a strong conflict is represented by $\bar{\mu}^o$.

expected payoff is $\mu_3^o(\theta_3 + \theta_2/\max\{\omega^o, \omega(\theta_B)\})$, and it is maximized when θ_B approaches $\theta_3 > 0$.

Optimal binary report. Because the board's action is binary (approve or reject), it is straightforward that the optimal CEO's report can be simplified into one with only two realizations. This aligns with the analysis in the main part of the paper with a continuous distribution where binary reports are also among the optimal reporting structures. Corollary OA-3 formally states this observation, and Figure 2 graphically illustrates the two equivalent characterizations of the optimal system in a simplex $\Delta(\Theta)$.

Corollary OA-3 (2D optimal CEO's signal)

- (i) For a strong conflict ($\theta_B > \theta_B^o$), the optimal distribution $\tilde{\Phi}$ is equivalent to a distribution Φ_{LD} over a subset $\{L, D\}$. The board approves the project after observing $r = D$ and rejects the project after observing $r = L$.
- (ii) For a weak conflict ($\theta_B < \theta_B^o$), the optimal distribution $\tilde{\Phi}$ is equivalent to a distribution Φ_{AH} over a subset $\{A, H\}$. The board approves the project after observing $r = H$ and rejects the project after observing $r = A$.

For a strong conflict, we coarsen the optimal 3D-report realizations with interim beliefs μ^A and μ^B at points A and B leading to rejection into a single realization with an interim belief

μ^L at point L on the line AB . Report $r = L$ implies that the project has either low or intermediate value with non-zero probability. For a weak conflict, we coarsen the optimal 3D-report realizations with interim beliefs μ^C and μ^D at points C and D , leading to approval into a single realization with an interim belief μ^H at point H on the line CD . Report $r = H$ implies that the project has either intermediate or high value with non-zero probability. Notice that $\mu^L \in NL^0$ and $\mu^H \in NL^1$; hence, following the coarse report L or H , the board indeed does not learn and selects the action outright.

Proofs

Proof of Proposition OA-1: For a given project category $t \in \{novel, routine\}$, let $W^t(\theta_B)$ be the shareholders' welfare when the board is of type θ_B and the CEO is willing to select project category t . Let $\bar{W}^t$ denote the maximal welfare for each project category, i.e., $\bar{W}^{routine} = W^{routine}(\theta_S)$ and $\bar{W}^{novel} = W^{novel}(\theta_H)$.

Part (i). We first prove that if $F(\theta) \leq G(\theta)$ for any θ (F and G are identical or F first-order stochastically dominates G), then $\bar{W}^{novel} \geq \bar{W}^{routine}$. We know that the shareholders' optimal decision cutoff for each project category is at θ_S . Applying this decision cutoff (i.e., for any $\theta < \theta_S$, the probability mass is shifted to an atom at θ_S where the shareholders' welfare is zero), we obtain adjusted distributions F^{adj} and G^{adj} that account for the rejections of projects with negative value. Namely, $F^{adj}(\theta) = G^{adj}(\theta) = 0$ if $\theta < \theta_S$, and $F^{adj}(\theta) = F(\theta)$ and $G^{adj}(\theta) = G(\theta)$ if $\theta \geq \theta_S$. Since the adjusted distributions preserve the first-order stochastic dominance (as well as identity), $F^{adj}(\theta) \leq G^{adj}(\theta)$ for any θ , novel projects are on average at least as profitable as routine projects when the shareholders can implement the optimal decision cutoff θ_S for any project category, $\bar{W}^{novel} = \int_{\theta_S}^{\theta_{max}} \theta f(\theta) d\theta \geq \int_{\theta_S}^{\theta_{max}} \theta g(\theta) d\theta = \bar{W}^{routine}$.

Thus, if $F(\theta) \leq G(\theta)$ for any θ , the shareholders primarily seek to obtain value $\bar{W}^{novel}$. This value is attainable by the shareholders. Namely, when the board is extreme, $\theta_B = \theta_H$, inequalities $W^{novel}(\theta_H) = \bar{W}^{novel} \geq \bar{W}^{routine} > W^{routine}(\theta_H)$ and $G(\theta_H) - F(R(\theta_H)) \geq F(\theta_H) - F(R(\theta_H)) > 0$ jointly imply $M(\theta_H) > 0$. (See also fn. 21.) In other words, the

CEO's and shareholders' preferences over project selection at $\theta_B = \theta_H$ are aligned, and this value is attainable.

Part (ii). By analogy to the claim above, if $F(\theta) \geq G(\theta)$ for any θ where $F(\theta) > G(\theta)$ for some θ (G first-order stochastically dominates F), then $\overline{W}^{novel} < \overline{W}^{routine}$. The shareholders seek to attain value $\overline{W}^{routine}$. If $M(\theta_S) < 0$, this value is attainable when $\theta_B = \theta_S$ because the CEO's and shareholders' preferences are aligned at $\theta_B = \theta_S$.

The value $\overline{W}^{routine}$ is not feasible when $M(\theta_S) \geq 0$; the CEO's and shareholders' preferences over project selection are not aligned at $\theta_B = \theta_S$. We proceed in several steps. First, by contradiction, we prove $\theta_B^* \neq \theta_S$. Suppose not, $\theta_B^* = \theta_S$. Since $M(\theta_S) \geq 0$, it means that a combination of a neutral board and a novel project is optimal for the shareholders.

- When $R(\theta_S) > \theta_C$ (strong conflict with a neutral board), $R(\theta_B)$ is increasing if $\theta_B \in [\theta_S, \theta_H]$ and therefore $W^{novel}(\theta_B)$ is also increasing. For the neutral board to be constrained-optimal for novel projects, we must have that $M(\theta_B) < 0$ if $\theta_B \in (\theta_S, \theta_H]$ (i.e., any increase in θ_B forces the CEO to select the routine project). But if $M(\theta_S) > 0$, this is impossible due to continuity of M function. If $M(\theta_S) = 0$, the neutral board is not constrained-optimal because by continuity of $W^{routine}$ function, there exists a board $\theta_B \in (\theta_S, \theta_S + \varepsilon)$ which selects a routine project such that $W^{routine}(\theta_B) > \overline{W}^{novel} > W^{novel}(\theta_S)$.
- When $R(\theta_S) = \theta_C$ (weak conflict with a neutral board), $R(\theta_B)$ is constant if $\theta_B \in [\theta_S, H^{-1}(\theta_C)]$. Therefore, on this interval, the CEO's expected payoff from a novel project is also constant, whereas her expected payoff from a routine project is decreasing. As a result, $M(\theta_B) > M(\theta_S) \geq 0$ if $\theta_B \in (\theta_S, H^{-1}(\theta_C)]$. By continuity of M function, we must therefore have $M(\theta_B) > 0$ also for $\theta_B \in [H^{-1}(\theta_C), H^{-1}(\theta_C) + \varepsilon)$ where $R(\theta_B)$ is increasing. But if $R(\theta_B)$ is increasing, $W^{novel}(\theta_B)$ is also increasing, and the shareholders will strictly prefer decisions on the novel project with a biased board $\theta_B \in (H^{-1}(\theta_C), H^{-1}(\theta_C) + \varepsilon)$ to decisions on the novel project with a neutral board $\theta_B = \theta_S$. This is a contradiction.

Second, we prove $\theta_B^* > \theta_C$ by ruling out $\theta_B^* = \theta_C$ and $\theta_B^* < \theta_C$. We start by observing that a board aligned with the CEO, $\theta_B = \theta_C$, convinces the CEO to select a routine project.

This is because $R(\theta_C) = \theta_C$ and therefore, by first-order stochastic dominance, $M(\theta_C) = W^{novel}(\theta_C) - W^{routine}(\theta_C) < 0$. At the same time, recall that $M(\theta_S) \geq 0$. By continuity of M function, $M(\theta_B) < 0$ also on the neighborhood $\theta_B \in (\theta_C, \theta_C + \varepsilon)$ for some small $\varepsilon > 0$. But shareholders' welfare, $W^{routine}(\theta_B)$, is increasing on the neighborhood and thus $\theta_B^* \neq \theta_C$.

Next, suppose $\theta_B^* < \theta_C$. If the CEO selects a routine, the outcome with θ_B^* is equivalent to the one with $\theta_B = \theta_C$ because the decision cutoff for routine projects is at $\max\{\theta_B^*, \theta_C\} = \theta_C$. If the CEO selects a novel project, θ_B^* is suboptimal for the shareholders because a higher welfare is achieved for $\theta_B = \theta_S$ where the CEO also selects a novel project.

Proof of Proposition OA-2: We proceed in two steps. In the first step, we take reporting strategy $\theta_R = R(\theta_B)$ and presentation strategy $(d_l, d_h) = (0, 1)$ as given and prove that only a babbling cheap talk equilibrium exists. In the second step, we prove that the CEO does not deviate from the reporting and presentation strategies, $R(\theta_B)$ and $(d_l, d_h) = (0, 1)$.

Step 1: We analyze only report-specific cheap talk for $r = h$, because the project is not presented for $r = l$. Conditional on high report, $\theta \in M_h \equiv [\theta_R, \theta_{max}]$. There exists a babbling equilibrium where the CEO sends a single message that θ belongs to M_h : in this case the project is approved. It is easy to see that there does not exist a two-message equilibrium. In such an equilibrium the interval M_h is partitioned into (M_h^-, M_h^+) . Because the board approves the project when the message is that $\theta \in M_h$, it also approves the project when the message is that $\theta \in M_h^+$. However, if the CEO observes $\theta \in M_h^-$, she communicates that $\theta \in M_h^+$ because she wants the project to be approved (see $\theta \geq \theta_R = R(\theta_B) \geq \theta_C$). Thus the communication degenerates into babbling.

Step 2: In the previous step we showed that when the reporting cutoff is $\theta_R = R(\theta_B)$, the cheap talk is uninformative and the board's information is characterized by a partition at θ_R . Is there any other partition that is better for the CEO from an ex ante perspective? By Lemma 4 of the main text, the partition at the cutoff $R(\theta_B)$ is the best for the CEO partition under which the board is willing to approve the project after message $r = h$. A finer partitioning cannot increase the CEO's expected payoff. To do so, such a partition must result in project approval for some $\theta \in (\theta_C, R(\theta_B))$, which is impossible since $\theta = R(\theta_B)$ is the lowest value of a project that the board approves in any partition.

Proof of Proposition OA-3: Consider the case where the constraint for the upper bound is loose, $L(\theta_C + \gamma) \geq \theta_S$. Then, a board with $\theta_B = \theta_H$ implements the shareholders' first best outcome. Now suppose that this constraint is binding, $L(\theta_C + \gamma) < \theta_S$. If the board is characterized by $\theta_B = \theta_H$, only an uninformative pooling equilibrium where the project is rejected can arise. The shareholders compare two outcomes: (i) set $\theta_B = H^{-1}(L(\theta_C + \gamma)) \in (\theta_S, \theta_H)$ and incentivize informative communication with the lowest possible overinvestment, or (ii) set any $\theta_B \in (H^{-1}(L(\theta_C + \gamma)), \theta_{max}]$ and acquiesce to project rejection after uninformative communication. By our tie-breaking assumption, $\theta_B = \theta_{max}$ in this case. The former cutoff means overinvestment and the latter cutoff means underinvestment (in fact, zero investment), because $H^{-1}(L(\theta_C + \gamma)) < \theta_H < \theta_{max}$ and any board with $\theta_B < \theta_H$ overinvests while any board with $\theta_B > \theta_H$ underinvests in novel projects.

When $\theta_S \leq \mathbb{E}[\theta]$, the shareholders prefer option (i) because project approval after $r = h$ generates a positive value, $\mathbb{E}[\theta | \theta \geq \theta_R] > \mathbb{E}[\theta] \geq \theta_S$. When $\theta_S \geq \mathbb{E}[\theta]$, the shareholders prefer option (i) only when $L(\theta_C + \gamma) > H(\theta_S)$. (Note that the shareholders of type $\theta_S > \mathbb{E}[\theta]$ earn a non-negative project value if $\theta_R \geq H(\theta_S)$.) In contrast, when $H(\theta_S) > L(\theta_C + \gamma)$, the shareholders prefer option (ii).

Proof of Lemma OA-1: By Lemma 7 in the main text, for a given κ , the CEO optimally sets $\theta_R \in [R(\theta_B), \hat{R}(\theta_B, \kappa)]$ and expects the decision cutoff $\theta_M = \hat{R}(\theta_B, \kappa)$. Take the intersection of the intervals $[R(\theta_B), \hat{R}(\theta_B, \kappa)]$ for all feasible levels of κ , i.e., the interval $[R(\theta_B), \hat{R}(\theta_B, \bar{\kappa})]$. For any θ_R from this interval, the board's decision cutoff is, for a given κ , $\theta_M = \hat{R}(\theta_B, \kappa)$. Recall that this is an identical decision cutoff that the CEO would implement if she knew κ in advance. Therefore, it is optimal for the CEO to set $\theta_R \in [R(\theta_B), \hat{R}(\theta_B, \bar{\kappa})]$; this achieves an equivalent outcome, as if the CEO had observed the realization of an unknown κ .

Proof of Lemma OA-2: For given κ , we know that the optimal board is $\hat{\theta}_H = \min\{\theta_H, \theta_S + \kappa\}$, which is increasing in κ . For an uncertain κ , the optimal board must be in between the least conservative optimal board (i.e., the optimal board for the lowest learning cost) and the most conservative optimal board (i.e., the optimal board for the highest learning cost), $\min\{\theta_H, \theta_S + \underline{\kappa}\} \leq \hat{\theta}_H \leq \min\{\theta_H, \theta_S + \bar{\kappa}\}$. Since $\underline{\kappa} > 0$, the optimal board is biased, $\hat{\theta}_H > \theta_S$.

Proof of Lemma OA-3: When characterizing the extreme points of NL^0 and NL^1 , we apply Proposition 2 in Caplin, Dean, and Leahy (2019). It characterizes the optimal strategy of a rationally inattentive receiver (i.e., board) that satisfies the martingale property. In particular, we use the fact that if the project decision is binary (approve or reject) and an interim belief is located in a learning region, the optimal board information acquisition strategy is a lottery over two final beliefs. The first final belief, μ^t , is in the border of NL^1 (a point t) and leads to the project approval, and the second final belief $\mu^{t'}$ is in the border of NL^0 (a point t') and leads to the project rejection. To derive these borderline posterior beliefs, we apply the Invariant Likelihood Ratio (ILR) Equations for Chosen Options' property. Its geometric interpretation in 2D-space is that the slopes of the board's net payoff function at points t and t' , $\sum_j \mu_j^t u_B(1, \theta_j) + \kappa H(\mu^t)$ and $\sum_j \mu_j^{t'} u_B(0, \theta_j) + \kappa H(\mu^{t'})$, must be identical. Intuitively, this property uses the fact that the board's net payoff function emerges as a result of concavification and, after concavification, the payoff function is by definition linear in the learning region and strictly concave in non-learning regions. Figure 5 in Caplin, Dean, and Leahy (2019) illustrates this property.

We begin with deriving the extreme points D and F on the BC -line where $\mu_1 = 0$ and $\mu_2 + \mu_3 = 1$. In the borderline point D , the board approves the project, and the board's net payoff function is $\mu_3^D(\theta_3 - \theta_B - \kappa \ln \mu_3^D) + (1 - \mu_3^D)(\theta_2 - \theta_B - \kappa \ln(1 - \mu_3^D))$; its slope in μ_3 -dimension is $(\theta_3 - \theta_B - \kappa \ln \mu_3^D) - (\theta_2 - \theta_B - \kappa \ln(1 - \mu_3^D))$. In the borderline point F , the board rejects the project, and the board's net payoff function is $-\mu_3^F \kappa \ln \mu_3^F - (1 - \mu_3^F) \kappa \ln(1 - \mu_3^F)$. Its slope in μ_3 -dimension is $-\kappa \ln \mu_3^F + \kappa \ln(1 - \mu_3^F)$. As in Caplin, Dean, and Leahy (2019), we use that the slopes are equal if the state-specific components are equal across states, $\theta_3 - \theta_B - \kappa \ln \mu_3^D = -\kappa \ln \mu_3^F$ and $\theta_2 - \theta_B - \kappa \ln(1 - \mu_3^D) = -\kappa \ln(1 - \mu_3^F)$. After rearranging, $e^{\frac{1}{\kappa}(\theta_3 - \theta_B)} = \frac{\mu_3^D}{\mu_3^F}$ and $e^{\frac{1}{\kappa}(\theta_2 - \theta_B)} = \frac{1 - \mu_3^D}{1 - \mu_3^F}$. The unique solution to this system of two linear equations is

$$(\mu_3^D, \mu_3^F) = \left(\frac{e^{\frac{1}{\kappa}(\theta_3 - \theta_2)} - e^{\frac{1}{\kappa}(\theta_3 - \theta_B)}}{e^{\frac{1}{\kappa}(\theta_3 - \theta_2)} - 1}, \frac{e^{\frac{1}{\kappa}(\theta_B - \theta_2)} - 1}{e^{\frac{1}{\kappa}(\theta_3 - \theta_2)} - 1} \right).$$

Note that for $\theta_B \in (\theta_2, \theta_3)$, it always holds that $\mu_3^D \in (0, 1)$ and $\mu_3^F \in (0, 1)$. That is, NL^1 is non-empty.

The extreme points E and G are on the AC -line where $\mu_2 = 0$ and $\mu_1 + \mu_3 = 1$. In the borderline point E , the board approves the project, and the board's net payoff function is $\mu_3^E(\theta_3 - \theta_B - \kappa \ln \mu_3^E) + (1 - \mu_3^E)(\theta_1 - \theta_B - \kappa \ln(1 - \mu_3^E))$. Its slope in μ_3 -dimension is $(\theta_3 - \theta_B - \kappa \ln \mu_3^E) - (\theta_1 - \theta_B - \kappa \ln(1 - \mu_3^E))$. In the borderline point G , the board rejects the project, and the board's net payoff function is $-\mu_3^G \kappa \ln \mu_3^G - (1 - \mu_3^G) \kappa \ln(1 - \mu_3^G)$. Its slope in μ_3 -dimension is $-\kappa \ln \mu_3^G + \kappa \ln(1 - \mu_3^G)$. Imposing that the state-specific components are equal across states and rearranging yields $e^{\frac{1}{\kappa}(\theta_3 - \theta_B)} = \frac{\mu_3^E}{\mu_3^G}$ and $e^{\frac{1}{\kappa}(\theta_1 - \theta_B)} = \frac{1 - \mu_3^E}{1 - \mu_3^G}$. The unique solution to this system of two linear equations is

$$(\mu_3^E, \mu_3^G) = \left(\frac{e^{\frac{1}{\kappa}(\theta_3 - \theta_1)} - e^{\frac{1}{\kappa}(\theta_3 - \theta_B)}}{e^{\frac{1}{\kappa}(\theta_3 - \theta_1)} - 1}, \frac{e^{\frac{1}{\kappa}(\theta_B - \theta_1)} - 1}{e^{\frac{1}{\kappa}(\theta_3 - \theta_1)} - 1} \right).$$

Note that for $\theta_B \in (\theta_2, \theta_3)$, it always holds that $\mu_3^E \in (0, 1)$ and $\mu_3^G \in (0, 1)$. That is, NL^0 is non-empty.

Proof of Lemma OA-4: The proof is by contradiction. Let $W(\mu) \equiv \mathbb{E}[u_C(\cdot) \mid r]$ be the expected payoff of the CEO at the interim stage for a given report realization (and the interim belief μ that it invokes). We refer to $W(\mu)$ as the CEO's "indirect value function."¹⁷ The optimal report is characterized by the distribution $\tilde{\phi}$ over the reduced set $\hat{\mathbb{S}}$, where

$$\tilde{\phi} = \left(\tilde{\phi}^r \right)_{r \in \hat{\mathbb{S}}} = \arg \max_{\phi \in \Delta(\hat{\mathbb{S}})} \sum_{r \in \hat{\mathbb{S}}} \phi^r W(\mu^r), \quad (\text{OA-2})$$

subject to the martingale property, $\mu^o = \sum_{r \in \hat{\mathbb{S}}} \phi^r \mu^r$.¹⁸ We show that any distribution with a positive ϕ^E , ϕ^F , or ϕ^G is not optimal.

- (i) Suppose $\tilde{\phi}^E > 0$: Construct a refined distribution $\hat{\phi}$ by redistributing all probability mass from point E to points A and C . The probability mass is divided into $(1 - \mu_3^E, \mu_3^E) = (\mu_1^E, \mu_3^E)$ shares such that the martingale property is satisfied, $(1 - \mu_3^E)\mu^A +$

¹⁷Because the report discourages additional information acquisition, the board's action at the interim stage (i.e., for a given belief μ) is a deterministic variable, characterized by a function, $\sigma : \Delta(\Theta) \rightarrow \{0, 1\}$. With this observation, we can easily express the expected payoff of the CEO at the interim stage: $W(\mu) = \sigma(\mu) \left(\sum_j \mu_j \theta_j - \theta_C \right)$.

¹⁸More precisely, the problem is characterized by 18 constraints, where 14 inequalities and one equality are due to the existence of a simplex, $0 \leq \phi^r \leq 1, r \in \hat{\mathbb{S}}, \sum_{r \in \hat{\mathbb{S}}} \phi^r = 1$, and three equalities are the martingale properties: $\mu_j^o = \sum_{r \in \hat{\mathbb{S}}} \phi^r \mu_j^r, j = 1, 2, 3$.

$\mu_3^E \mu^C = \mu^E$. That is, $\hat{\phi}_A = \tilde{\phi}^A + \mu_1^E \tilde{\phi}^E < 1$, $\hat{\phi}_C = \tilde{\phi}^C + \mu_3^E \tilde{\phi}^E < 1$, $\hat{\phi}_E = 0$, and $\hat{\phi} = \tilde{\phi}$ otherwise. The redistribution increases the CEO's objective,

$$\sum_{r \in \hat{\mathbb{S}}} \hat{\phi} W(\mu) - \sum_{r \in \hat{\mathbb{S}}} \tilde{\phi} W(\mu) = \tilde{\phi}^E [\mu_1^E W(\mu^A) + \mu_3^E W(\mu^C) - W(\mu^E)] = -\tilde{\phi}^E \mu_1^E W(\mu^A) > 0,$$

which follows from $\mu_1^E > 0$, $\tilde{\phi}^E > 0$, and $W(\mu^A) = \theta_1 - \theta_C < 0$.

- (ii) Suppose $\tilde{\phi}^F > 0$: By analogy, construct a refined distribution $\hat{\phi}$ by redistributing all probability mass from point F to points B and C . The probability mass is divided into $(1 - \mu_3^F, \mu_3^F) = (\mu_2^F, \mu_3^F)$ shares such that the martingale property is satisfied, $(1 - \mu_3^F) \mu^B + \mu_3^F \mu^C = \mu^F$. That is, $\hat{\phi}_B = \tilde{\phi}^B + \mu_2^F \tilde{\phi}^F < 1$, $\hat{\phi}_C = \tilde{\phi}^C + \mu_3^F \tilde{\phi}^F < 1$, $\hat{\phi}_F = 0$, and $\hat{\phi} = \tilde{\phi}$ otherwise. The redistribution increases the CEO's objective,

$$\sum_{r \in \hat{\mathbb{S}}} \hat{\phi} W(\mu) - \sum_{r \in \hat{\mathbb{S}}} \tilde{\phi} W(\mu) = \tilde{\phi}^F [\mu_2^F W(\mu^B) + \mu_3^F W(\mu^C) - W(\mu^F)] = \tilde{\phi}^F \mu_3^F W(\mu^C) > 0,$$

which follows from $\mu_3^F > 0$, $\tilde{\phi}^F > 0$, and $W(\mu^C) = \theta_3 - \theta_C > 0$.

- (iii) Suppose $\tilde{\phi}^G > 0$: As in the case of point E , construct a refined distribution $\hat{\phi}$ by redistributing all probability mass from point G to points A and C . The probability mass is divided into $(1 - \mu_3^G, \mu_3^G) = (\mu_1^G, \mu_3^G)$ shares such that the martingale property is satisfied, $(1 - \mu_3^G) \mu^A + \mu_3^G \mu^C = \mu^G$. That is, $\hat{\phi}_A = \tilde{\phi}^A + \mu_1^G \tilde{\phi}^G < 1$, $\hat{\phi}_C = \tilde{\phi}^C + \mu_3^G \tilde{\phi}^G < 1$, $\hat{\phi}_G = 0$, and $\hat{\phi} = \tilde{\phi}$ otherwise. The redistribution increases the CEO's objective,

$$\sum_{r \in \hat{\mathbb{S}}} \hat{\phi} W(\mu) - \sum_{r \in \hat{\mathbb{S}}} \tilde{\phi} W(\mu) = \tilde{\phi}^G [\mu_1^G W(\mu^A) + \mu_3^G W(\mu^C) - W(\mu^G)] = \tilde{\phi}^G \mu_3^G W(\mu^C) > 0,$$

which follows from $\mu_3^G > 0$, $\tilde{\phi}^G > 0$, and $W(\mu^C) = \theta_3 - \theta_C > 0$.

Proof of Proposition OA-4: Let, as in the proof of Lemma OA-4, $W(\mu) \equiv \mathbb{E}[u_C(\cdot)|r]$ be the CEO's indirect value function. We analyze whether an alternative convex hull exists such that the prior lies in the hull and a signal can be constructed from extreme points of the hull. By Carathéodory's theorem, concavification over a two-dimensional simplex is based on

at most three linearly biased points. In our case, convex hulls created by triplets of linearly biased points are $\{ABD, ACD, ABC\}$. In addition, we have convex hulls constructed by pairs of points, i.e., lines. We proceed in two steps. First, we suppose that the prior is not on any line between points $\{A, B, C, D\}$, which is equivalent to being in the interior of Δ_{ABD} or in the interior of Δ_{ACD} . This eliminates convex hulls constructed by pairs of points. Second, we analyze the boundaries of Δ_{ABD} and Δ_{ACD} .

- For any prior in the interiors, $\mu_0 \in \text{int}(\Delta_{ABD}) \cup \text{int}(\Delta_{ACD})$, the only alternative convex hull is Δ_{ABC} . Replacing $\tilde{\phi}$ (where $\tilde{\phi} = \phi_{ABD}$ or $\tilde{\phi} = \phi_{ACD}$) by ϕ_{ABC} is equivalent to redistributing all probability mass $\tilde{\phi}^D$ from point D to points B and C . The probability mass is divided into $(1 - \mu_3^D, \mu_3^D) = (\mu_2^D, \mu_3^D)$ shares such that the martingale property is satisfied, $(1 - \mu_3^D)\mu^B + \mu_3^D\mu^C = \mu^D$. That is, $\phi_{ABC}^B = \tilde{\phi}^B + \mu_2^D\tilde{\phi}^D < 1$, $\phi_{ABC}^C = \tilde{\phi}^C + \mu_3^D\tilde{\phi}^D < 1$, $\phi_{ABC}^D = 0$, and $\phi_{ABC} = \tilde{\phi}$ otherwise. This redistribution decreases the CEO's objective,

$$\begin{aligned} \sum_{r \in \hat{\mathbb{S}}} \phi_{ABC} W(\mu) - \sum_{r \in \hat{\mathbb{S}}} \tilde{\phi} W(\mu) &= \tilde{\phi}^D [\mu_2^D W(\mu^B) + \mu_3^D W(\mu^C) - W(\mu^D)] \\ &= -\tilde{\phi}^D (1 - \mu_3^D) W(\mu^B) < 0, \end{aligned}$$

because $\mu_2^D > 0$ and $\mu_3^D > 0$ (hence, $\tilde{\phi}^D > 0$), and $W(\mu^B) = \theta_2 - \theta_C > 0$.

- Consider boundaries of the simplex. For $\mu_1^D = 0$, we replicate the argument from above; redistribution of the probability mass from point D to points B and C decreases the CEO's objective as $W(\mu^D) > \mu_2^D W(\mu^B) + \mu_3^D W(\mu^C)$. For $\mu_2^D = 0$, all feasible signal distributions based on $\{A, B, C, D\}$ are exactly equivalent, $(\phi^A, \phi^B, \phi^C, \phi^D) = (\mu_1^D, 0, \mu_3^D, 0)$. The same holds for $\mu_3^D = 0$, where $(\phi^A, \phi^B, \phi^C, \phi^D) = (\mu_1^D, \mu_2^D, 0, 0)$.
- The remaining case is when μ_0 is located on AD line but not on the boundary. Then, $\frac{\mu_2^D}{1 - \mu_3^D} = \frac{\mu_3^D}{\mu_3^D}$, and $\phi_{ABD} = \phi_{ACD}$. No other distribution based on $\{A, B, C, D\}$ is feasible.

To sum up, none of the alternative concavifications over $\hat{\mathbb{S}}$ is a solution to the CEO's linear programming problem in equation (OA-2).

Proof of Corollary OA-2: Follows directly from Proposition OA-4 and is omitted.

Proof of Corollary OA-3: For a strong conflict ($\mu_0 \in \Delta_{ABD}$) we can coarsen the report realizations with interim beliefs μ^A and μ^B at points A and B leading to rejection, into a single realization with interim belief μ^L at point L , where

$$\mu^L = \frac{\tilde{\phi}^A}{\tilde{\phi}^A + \tilde{\phi}^B} \mu^A + \frac{\tilde{\phi}^B}{\tilde{\phi}^A + \tilde{\phi}^B} \mu^B.$$

The optimal 3D-distribution then is equivalent to a 2D-distribution ϕ_{LD} over a subset $\{L, D\}$ where

$$(\phi_{LD}^L, \phi_{LD}^D) \equiv (\phi_{ABD}^A + \phi_{ABD}^B, \phi_{ABD}^D).$$

For a weak conflict ($\mu_0 \in \Delta_{ACD}$) we can coarsen the report realizations with interim beliefs μ^C and μ^D at points C and D leading to approval, into a single realization with interim belief μ^H at point H , where

$$\mu^H = \frac{\tilde{\phi}^C}{\tilde{\phi}^C + \tilde{\phi}^D} \mu^C + \frac{\tilde{\phi}^D}{\tilde{\phi}^C + \tilde{\phi}^D} \mu^D.$$

The optimal 3D-distribution is equivalent to a 2D-distribution ϕ_{AH} over a subset $\{A, H\}$, where

$$(\phi_{AH}^A, \phi_{AH}^H) \equiv (\phi_{ACD}^A, \phi_{ACD}^C + \phi_{ACD}^D).$$

References

- Caplin, A., M. Dean, and J. Leahy. 2019. Rational inattention, optimal consideration sets, and stochastic choice. *The Review of Economic Studies* 86 (3): 1061–94.
- Friedman, H., J. Hughes, and B. Michaeli. 2020. Optimal reporting when additional information might arrive. *Journal of Accounting and Economics* 69 (2–3): 101276.
- Matysková, L., and A. Montes 2023. Bayesian persuasion with costly information acquisition. *Journal of Economic Theory* 211: 105678.