Innovation Races with Public Benchmarks:*

Theory and Evidence from AI

Edward Brookman

June 22, 2026

Abstract

Benchmark leaderboards have become a central institution for evaluating AI models. This paper studies the competitive environment created by such leaderboards. I first exploit a methodological change to an LLM leaderboard to provide causal evidence that benchmark scores drive model demand. Using this evidence, I develop an innovation race in which firms compete for capital based on the benchmark performance of their models. Firms can improve their scores either by investing in persistent quality improvements or by engaging in rent-seeking 'benchmark hacking'. The model exhibits an 'efficiency curse' in which the most productive firms are the weakest competitors on the leaderboard. In the short run, the leaderboard misallocates capital to inefficient firms, allowing them to keep pace with their rivals. In the long run, inefficient firms either 'run out of steam' or remain the leader indefinitely, depending on the growth rate of the industry as a whole. I show that benchmark hacking degrades the information content of leaderboards but does not eliminate it. In the absence of better institutions, rational consumers rely on imperfect scores, consistent with my empirical evidence. The framework yields concrete guidance for benchmark design in competitive industries.

*The author has no sources of research support, financial relationships, or other potential conflicts of interest to disclose.

1 Introduction

Leaderboards are present in many domains of life. Students are ranked by test scores, salespeople by revenue, and universities by research output. In each case a proxy for a complex underlying quality is made public and the resulting leaderboard shapes behaviour on both sides of the market; participants adjust effort in anticipation of where they will place, and outsiders use the leaderboard to allocate attention, patronage, and capital.

A leaderboard is, in this sense, a contest. By making relative performance salient, it converts what might otherwise be a collection of independent optimisation problems into a strategic game among ranked participants. The economic literature on contests has long studied such games and generally emphasizes their capacity to stimulate effort, innovation, and competition [Corchón, 2007, Halac et al., 2017, Lemus and Marshall, 2021]. In many important settings, however, leaderboards are dynamic and administer prizes *during* a contest. A high position on the leaderboard attracts funding, attention, contracts, and customers, which in turn determine the resources available for future competition. For example, consider the race to build powerful artificial intelligence, which serves as our motivating context. AI labs release models, which are tested and ranked on public leaderboards.[1] Their ranking determines the number of users and investors they attract, subsequently influencing the quality of the next model they can release. Here, the presence of a public leaderboard serves as an allocation mechanism that may or may not provide a good environment for innovation.

This paper studies the incentives created by such allocation rules when leaderboard scores can be artificially inflated at a cost. This is true in many settings, typically by overfitting to the measured metric. In particular, it is known that AI companies can inflate benchmark test performance without improving the underlying model. This behaviour is often termed '*Benchmaxxing*', '*Benchmark hacking*' or '*SOTA-Chasing*' [2]. For example, Pfister and Jud [2025] estimate that OpenAI spent on the order of hundreds of thousands of dollars in compute costs to achieve a high score on the ARC-AGI benchmark; Singh et al. [2025] report that Meta created 27 versions of Llama-4 before selecting the one that achieved the highest score on a public leaderboard.

[1] Popular such leaderboards include LMArena: https://arena.ai/leaderboard/ and Artificial Analysis: https://artificialanalysis.ai/leaderboards/models

[2] SOTA stands for 'State of the art'

This paper develops a tractable model of innovation and competition between two firms in which a leaderboard shapes the path of innovation. Firms are rewarded ultimately for performing genuine research but may divert resources toward temporarily inflating their scores on the leaderboard. Such 'score hacking' is rent-seeking, but steals users and investment that may fund next period's decisions. I ask, to what extent does the leaderboard; direct capital to its most efficient use, reward the most productive firms and stimulate innovation?

I find that, contrary to intuition, leaderboards can systematically favour less efficient firms and provide bad welfare outcomes. The driving mechanism is that, when leaderboards can be gamed, firms must 'hack' to secure future funding. However, the opportunity cost of doing so is highest exactly for the firms that are best at genuine research. In other words, the leaderboards introduce a multi-tasking problem that distract efficient firms from their comparative advantage. Taken to its extreme, I find that a firm strictly dominated on research efficiency, can achieve a higher payoff in the game.

That being said, the leaderboard can prove useful as the game progresses. As the underlying technology of inefficient firms begin to fall behind, scores may separate despite attempts to inflate it. Once this happens, the leaderboard directs capital toward productive competitors, improving total innovation. This separation is not guaranteed however. If the industry (total users and investment) is growing sufficiently fast, an inefficient firm can maintain the top spot on the leaderboard indefinitely. The leaderboard both encourages hacked scores, and rewards the hackers with the resources to continue hacking. I use the model to comment on design choices of leaderboards and how they shape the incentives to manipulate them. I apply this to the context of benchmark design, which is an active area of discussion in computer science and public rhetoric.

To test the assumptions of my theoretical model, I also conduct an event study using data from the open-source AI platform 'Hugging Face'. I exploit a methodology change that reshuffled rankings on a leaderboard whilst keeping the participants the same. I find evidence that leaderboards reallocate user adoption, independent of model quality.

Related Literature. This paper contributes primarily to the literature on dynamic contests and innovation races, and more broadly to work on performance measurement when incentives are attached to manipulable public metrics.

Dynamic contest theory studies how current standings shape later incentives through timing, feedback, and prize design. Yildirim [2005] studies multi-round contests in which players can revise effort after observing earlier play, while Aoyagi [2010] and Hinnosaar [2024] analyse how interim disclosure and sequencing affect behaviour in dynamic tournaments and sequential contests. In innovation settings, designers may use intermediate prizes, disclosure rules, or history-dependent rewards to influence experimentation and completion; see Halac et al. [2017], Benkert and Letina [2020], and Lemus and Marshall [2021]. The main concern in that literature is how a principal should disclose information about the other players to maximise total effort and experimentation. Instead of focusing on the impact of feedback, I study a setting in which a flow of resources is directly being allocated throughout the duration of a contest.

A complementary literature studies how performance metrics change the behaviour they are meant to measure. Holmstrom and Milgrom [1991] show that when agents perform multiple tasks, strong incentives on a measurable activity crowd effort away from other valuable but harder-to-measure tasks. Related work in contract design emphasises that a performance measure is useful only insofar as it is aligned with the principal's true objective, and misaligned measures invite gaming rather than production. Baker et al. [1994] explain why organisations often supplement hard metrics with subjective evaluation when objective measures are incomplete or gameable. Dranove et al. [2003] provide a canonical empirical example: cardiac-surgery report cards induced providers to avoid high-risk patients, improving measured performance while worsening outcomes for the sickest patients. Leaderboards are a particularly stark version of this problem because they compress performance into a public ranking that directly governs attention, users, and capital. Relative to the multitasking and performance-measurement literature, the novel margin is dynamic selection across firms: the leaderboard distorts not only effort within firms, but also which firms receive the resources to innovate next.

The closest papers in spirit are dynamic R&D race models such as Reinganum [1982] and Harris and Vickers [1987], which study heterogeneous firms in a dynamic process of capability accumulation. Structurally, however, the model is closer to frameworks studied in political economy, such as Meirowitz [2008], who analyse an electoral contest in which candidates choose costly campaign effort under asymmetries in voter support and campaign technology. Relative to the R&D race literature, my model inserts a manipulable public leaderboard between technological capability and

the resources that finance subsequent innovation. Future capital is allocated not from the underlying state of the race itself, but from a public and gameable performance signal. The theoretical form also differs slightly: rather than a classic patent-race or stopping-race formulation, I study a finite-horizon dynamic allocation game with an evolving state variable.

Empirical work also studies how public leaderboards shape behaviour in practice. In Kaggle prediction contests, Lemus and Marshall [2021] combine observational data with a randomized intervention and find that real-time leaderboards improve contest outcomes on average by changing competitors' stopping and effort decisions. Hudja et al. [2025] show in a sequential-sampling experiment that public leaderboard feedback can instead reduce search effort and lower the expected quality of the winning entry, highlighting that the effect of feedback depends on contest structure. My empirical evidence complements this literature by shifting attention from participant effort to audience allocation: exploiting an exogenous reshuffling of the Hugging Face Open LLM Leaderboard, I provide field evidence that changes in rank reallocate user attention across models even when the underlying models are unchanged. This isolates the allocative role of rankings and speaks directly to how leaderboards operate on real platforms.

A large motivator for this research was to use the theoretical tools of economics to contribute to the emerging field of benchmark design in AI and computer science [Ott et al., 2022, Raji et al., 2021]. There is ongoing debate as to whether benchmarks, and the leaderboards they generate, encourage breakthroughs such as the 'DeepSeek moment'[3], or instead provide gameable measures that lead to distorted and anti-competitive effects [Alzahrani et al., 2024, Singh et al., 2025]. My work provides a game-theoretic framework to organise these observations and to identify conditions under which the positive and negative effects are most likely to arise. Very recently, other authors have begun work in this area. Chen et al. [2026] model benchmarking as a Stackelberg game in which a benchmark designer chooses an evaluation protocol and model developers respond by allocating post-training effort across benchmarks; their focus is on whether different evaluation rules induce equilibrium rankings that recover latent model quality, and on how interventions such as tune-before-test can eliminate benchmaxxing. Erlei et al. [2026], by contrast, study the downstream adoption side: in a 'market for lemons' environment for AI systems, they show how information

[3]The 'DeepSeek moment' refers to the January 2025 release of DeepSeek-R1, which topped leaderboards despite coming from a relatively unknown company.

asymmetries and disclosure design shape user adoption and market efficiency. My paper differs substantially from both by embedding flawed benchmarks within a dynamic setting. The central question is not only whether rankings reflect latent quality, or whether users adopt AI under asymmetric information, but how strategic score manipulation feeds back into future resources, competition, and welfare over time. In that sense, the paper offers a theory of benchmark design in which leaderboards affect the path of innovation itself.

The remainder of the paper is organised as follows. Section 2 presents the motivating empirical evidence. Section 3 introduces the model. Section 4 characterises the equilibria and interprets. Section 5 conducts the welfare analysis and discusses implications for benchmark design. Section 6 extends the model to consider sequential updating of the leaderboard. Section 7 gives a rational-demand microfoundation for leaderboard-based allocation. Section 8 concludes.

2 Motivating Evidence

The theoretical analysis conducted in the following sections rely on the assumption that high scores on the leaderboard attract more resources, and thus are worth fighting over. In this section, we test this assumption in the context of AI development using data from Hugging face. Hugging Face is the central distribution hub for open-source machine-learning models: researchers and firms publish model weights there, and users discover, download, and fine-tune them through the platform. Among the most widely referenced summaries of model performance on the platform is the Open LLM Leaderboard, a public ranking that since 2023 has scored open-weight language models on a fixed battery of benchmarks and displayed them in rank order on a platform homepage.

The original leaderboard (henceforth *v1*) averaged raw scores across six benchmarks (ARC-Challenge, HellaSwag, MMLU, TruthfulQA, Winogrande, and GSM8K). By mid 2024, Hugging Face judged that this design had degraded in several respects. Several of the benchmarks had become saturated, with frontier models clustered near the ceiling and therefore hard to distinguish; others were suspected of contamination, having plausibly leaked into the pretraining corpora of models subsequently evaluated on them.

On June 27, 2024, Hugging Face replaced v1 with a redesigned *v2* [4]. The benchmark set

[4]Hugging Face Open LLM Leaderboard announcement: https://huggingface.co/spaces/open-llm-leaderboard/blog.

was replaced with a harder and less-saturated battery and certain tweaks were made to their methodology. The changes caused a reshuffling of models on the leaderboard, while keeping the underlying models fixed. We use this reshuffling as exogenous variation in leaderboard rank to isolate the causal effect on user adoption.

The rankings are reconstructed directly from the datasets backing the two leaderboards, so the ranks correspond exactly to the ordering a user would have seen on the site. Monthly downloads are measured at three pre-shock dates (approximately October 2023, March, 2024, and April, 2024) and seven post-shock dates running from August 8, 2024 through June 2025. Pre-shock counts come from existing research snapshots and a broad catalogue pull of Hugging Face models; post-shock counts come from pinned snapshots of Hugging Face's public hub statistics. Appendix B.1 and Appendix B.2 give the full data-construction details.

Let $\Delta\text{Rank}_i \equiv \text{rank}_i^{v1} - \text{rank}_i^{v2}$, so positive values indicate an improvement in leaderboard position between v1 and v2. The identification argument is a parallel-trends assumption: in the absence of the June 27 reshuffle, models with different realised values of ΔRank_i would have followed common trends in log downloads. I test this directly using a two-way fixed-effects panel event study,

$$\ln(1 + \text{dl}_{it}) \;=\; \alpha_i \;+\; \gamma_t \;+\; \sum_{\tau \neq \text{Apr 2024}} \beta_\tau \cdot \Delta\text{Rank}_i \cdot \mathbf{1}\{t = \tau\} \;+\; \varepsilon_{it}, \tag{1}$$

where α_i is a model fixed effect and γ_t a period fixed effect. The coefficients β_τ trace the effect of a one-rank improvement at each horizon, relative to the April 2024 reference period. Standard errors are clustered at the model level.

Figure 1 plots the estimated β_τ. Two features support the identifying assumption. First, the two estimated pre-shock event-time coefficients (October 2023 and March 2024, both measured relative to the omitted April 2024 reference) are economically and statistically indistinguishable from zero, and a joint Wald test of $\beta_{\text{Oct 2023}} = \beta_{\text{Mar 2024}} = 0$ fails to reject at conventional levels ($F = 0.59$, $p = 0.56$). Over an eight-month pre-window, models that would later improve in rank show no differential download growth. Second, the post-shock coefficients are positive at every horizon and generally become larger at longer horizons.

Pooling all post-shock periods into a single indicator interacted with ΔRank_i, the average effect is $\hat{\beta}_{\text{post}} = +0.0082$ (SE $= 0.0031$, $p = 0.008$). Collapsing the panel to a two-period cross-sectional

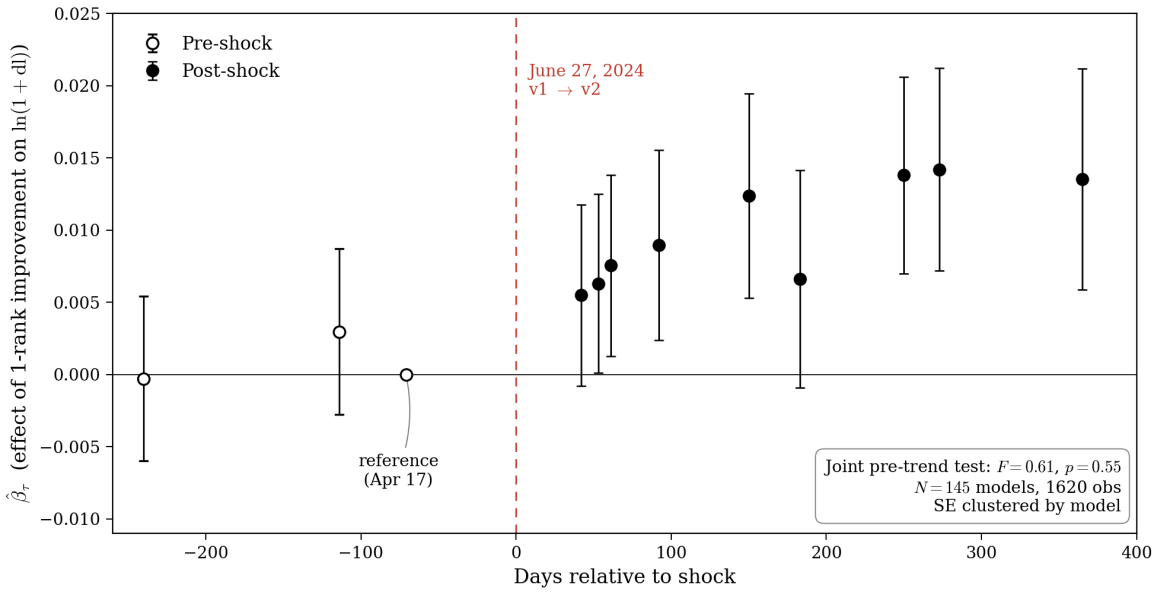


Figure 1: Event-study coefficients from equation (1). Open circles are pre-shock event times (October 2023, March 2024, and the omitted April 17, 2024 reference); filled circles are post-shock. Bars are 95% confidence intervals from standard errors clustered at the model level. The dashed red line marks the June 27, 2024 leaderboard replacement. Sample: 144 models. Pre-shock coefficients are jointly insignificant ($F = 0.59$, $p = 0.56$).

first-difference around the shock, using an average of the December 2024, March 2025, and June 2025 snapshots as the post-period outcome, gives a similar estimate, $\hat{\beta} = +0.0101$ (SE $= 0.0033$, $p = 0.002$, $N = 116$). The first-difference coefficient implies that a one-hundred-rank improvement in leaderboard position is associated with roughly a one log-point increase in monthly downloads, approximately a 2.75x increase in monthly downloads. Appendix B.3 and Appendix B.4 present further robustness.

The analysis above establishes that leaderboards can reallocate resources, in this case user downloads, as they change over time. The remainder of this paper considers the implications of this mechanism on the strategic behaviour of firms within an R&D race.

3 Model

3.1 Firms and quality accumulation.

Two firms $i \in \{1, 2\}$ compete to improve the quality of their technology, which begins the game at $Q_{i0} = 0$. Time is discrete and indexed by $t \in \{1, \ldots, T\}$ with $T \geq 2$. In each period, firm i

allocates their available budget B_{it} between research $r_{it} \geq 0$ and hacking $h_{it} \geq 0$. Investing in research improves the firm's technology according to $Q_{it} = Q_{i,t-1} + \alpha_i r_{it}$ where α_i represents firm i's research efficiency. The budget constraint requires $r_{it} + h_{it} = B_{it}$ and there is no saving between periods.

At the end of the game, firm i earns a payoff equivalent to the final quality of their technology: $U_i = Q_{iT}$.

3.2 Leaderboard scores and hacking

After each intermediate period $t \in \{1, \ldots, T-1\}$, firms are evaluated on a benchmark test and awarded a leaderboard score s_{it}. The firm's score reflects a combination of the true state of their technology and their hacking efforts. Specifically, the score of firm i in period t is given by

$$s_{it} = \theta Q_{it} + h_{it}. \tag{2}$$

The parameter $\theta > 0$ represents how strongly the leaderboard reflects true quality, while hacking is normalised to enter one-for-one into leaderboard performance. We assume $\theta \alpha_i < 1$ for both firms, so that reallocating one dollar from research to hacking raises the current leaderboard score.

In the final period ($t = T$), there is no further evaluation and thus both firms invest their entire terminal budget into research ($r_{iT} = B_{iT}$).

3.3 Capital allocation

How a firm scores on the leaderboard determines how much capital they attract for the next period. The demand microfoundation in Section 7 shows how Bayesian consumers can generate a positive demand response to a gameable score, and how this response aggregates to the Tullock allocation rule used below. In the dynamic race, after each contest period $t \in \{1, \ldots, T-1\}$, a fixed prize $M_t > 0$ is allocated between the two firms to form their budgets for the subsequent period. Firm i's budget in period $t+1$ is proportional to its share of the aggregate score in period t:

$$B_{i,t+1} = M_t \frac{s_{it}}{s_{it} + s_{jt}}. \tag{3}$$

Both firms enter the first period with an exogenously given initial budget $B_{i1} > 0$.

3.4 Information and strategies

The game is one of complete information: all parameters ($\{\alpha_i\}$, θ, $\{M_t\}_{t=1}^{T-1}$) and initial budgets are common knowledge, and actions and scores are publicly observed at the end of each period. Within period t, firms choose (r_{it}, h_{it}) simultaneously. A strategy σ_i for firm i specifies, at every history, a feasible choice (r_{it}, h_{it}) with $r_{it} + h_{it} = B_{it}$.

Definition 1 (Subgame-perfect equilibrium)**.** A strategy profile $\sigma^* = (\sigma_1^*, \sigma_2^*)$ is a *subgame-perfect equilibrium* if, at every history h^t and for each firm i,

$$U_i(\sigma_i^*, \sigma_j^* \mid h^t) \;\geq\; U_i(\sigma_i, \sigma_j^* \mid h^t) \qquad \text{for all strategies } \sigma_i,$$

where $U_i(\sigma \mid h^t)$ denotes firm i's continuation payoff from history h^t onward.

4 Equilibrium

Before turning to the formal characterization, it is useful to summarize the firm's problem in words. Firms ultimately care about terminal quality, so only increase their payoff when conducting research. In the intermediate contest periods, however, they are tempted to divert resources toward hacking. Hacking is 'cheap' contest effort but is dissipated after each period, whereas genuine research is 'expensive' but durable. The firm's problem is therefore dynamic: hacking steals funding from rivals but means that you enter next period behind in terms of true technology.

This dynamic trade-off collapses to a simple one-dimensional object. Define the *hacking incentive* as

$$\gamma_i \;\equiv\; 1 - \theta\alpha_i \;\in\; (0, 1),$$

γ_i represents the net score gain to firm i from reallocating one unit of budget from research to hacking. An extra unit of budget put into hacking raises the current score by one through direct manipulation but forfeits $\theta\alpha_i$ points because the true quality gain is forgone. We assume $\gamma_i > 0$ for both firms so that hacking is the strictly cheaper way to raise your score.

Taking firm j's strategy as given, firm i chooses $\{(r_{it}, h_{it})\}_{t=1}^{T}$ to solve

$$\max_{\{(r_{it},h_{it})\}} Q_{iT} \qquad \text{s.t.} \qquad r_{it} + h_{it} = B_{it}, \quad r_{it}, h_{it} \geq 0,$$

with $Q_{it} = Q_{i,t-1} + \alpha_i r_{it}$ and $B_{i,t+1} = M_t\, s_{it}/(s_{it} + s_{jt})$ at scores $s_{it} = \theta Q_{it} + h_{it}$.

Definition 2 (Feasible score interval)**.** For a given inherited quality and budget pair $(Q_{i,t-1}, B_{it})$, firm i can choose r_{it} to induce any score in the interval

$$[\underline{s}_{it}, \overline{s}_{it}] \;\equiv\; [\theta Q_{i,t-1} + \theta\alpha_i B_{it},\; \theta Q_{i,t-1} + B_{it}],$$

where $\underline{s}_{it}$ is the score under full research ($r_{it} = B_{it}$) and $\overline{s}_{it}$ is the score under full hacking ($r_{it} = 0$).

Throughout this section, we focus first on games satisfying the interiority condition. That is, prizes are neither too small nor too large such that Nash strategies prescribe both firms to choose a non-zero quantity of research and hacking in every contest period. Equivalently, along the equilibrium path the parameters must satisfy:

$$\underline{s}_{it} \;<\; \frac{\gamma_i^2 \gamma_j}{(\gamma_i + \gamma_j)^2} M_t \;<\; \overline{s}_{it} \qquad \forall\, i \in \{1, 2\},\; t < T.$$

This is the parameter region in which both firms still face a genuine trade-off between improving the technology and inflating the leaderboard score, rather than a corner in which one instrument is no longer used. This assumption is relaxed later.

The game is finite and thus solved by backward induction. In the terminal period, there is no longer a prize to fight over so both firms conduct research only: $r_{iT}^{*} = B_{iT}$. Consequently, a unit of budget entering T is worth α_i in terminal quality. In earlier contest periods, genuine research plays a dual role: it raises terminal quality directly and it raises the minimum score $\underline{s}_{it}$ in every subsequent evaluation. At the margin, both direct research and additional future budget contribute to terminal quality at rate α_i, so α_i scales both channels symmetrically and drops out of the first-order condition. What remains is the net score advantage from reallocating a unit of budget toward hacking, namely γ_i. Consequently, each contest subgame behaves like a Tullock contest in which firm i's effective strength is its hacking incentive.

Theorem 1 (Interior equilibrium characterization)**.** *Suppose the primitives and initial state satisfy the interiority condition in every continuation subgame. Then the game has a unique subgame-perfect equilibrium. At every contest date $t < T$, firm i chooses*

$$r_{it}^* = \frac{\theta Q_{i,t-1} + B_{it} - s_{it}^*}{\gamma_i}, \qquad h_{it}^* = B_{it} - r_{it}^*, \tag{4}$$

where the leaderboard scores and next-period budgets are

$$s_{it}^*, = \frac{\gamma_i^2 \gamma_j}{(\gamma_i + \gamma_j)^2} M_t, \qquad B_{i,t+1}^* = \frac{\gamma_i}{\gamma_i + \gamma_j} M_t \tag{5}$$

The resulting equilibrium payoff for firm i is

$$U_i^* \;=\; \frac{\alpha_i}{\gamma_i^{T-1}}\, B_{i1} \;+\; \frac{\alpha_i\, \gamma_i^2}{(\gamma_i + \gamma_j)^2} \sum_{t=1}^{T-1} \frac{M_t}{\gamma_i^{T-1-t}}. \tag{6}$$

The proof, in Appendix A.1, proceeds by backward induction on continuation values. The key observation is that the continuation value at the start of period t depends on the state only through the composite $Q_{i,t-1} + \alpha_i B_{it}$, and does so linearly:

$$V_i^t(Q, B) \;=\; \frac{Q_{i,t-1} + \alpha_i B_{it}}{\gamma_i^{T-t}} + D_i^t,$$

where D_i^t collects prize-dependent constants. In other words, accumulated quality and current budget are perfect substitutes in continuation value: both produce terminal quality at rate α_i. The first-order condition in any contest period therefore reduces to

$$M_t\, \gamma_i\, \frac{s_{jt}}{(s_{it} + s_{jt})^2} \;=\; 1,$$

which is the standard Tullock condition with each firm's effective strength equal to its hacking incentive (γ_i and γ_j). This is fixed and time invariant, so the same structure appears in every period, scaled only by the prize M_t.

Symmetric baseline. Before exploring the implications of firm heterogeneity, it is instructive to consider the symmetric baseline in which both firms possess identical research efficiency, $\alpha_1 =$

$\alpha_2 \equiv \alpha$, and identical initial endowments. Let $\gamma \equiv 1 - \theta\alpha$ denote the common hacking incentive. Naturally, each firm captures exactly half of the available prize in every contest period:

$$B^*_{i,t+1} = \frac{M_t}{2}, \qquad t \in \{1, \ldots, T-1\}.$$

Likewise, each firm posts the same leaderboard score in every contest period, $s^*_{it} = \gamma M_t/4$. Despite this balanced allocation, both firms still engage in strictly positive leaderboard hacking in equilibrium. This is the familiar logic of rent dissipation: even identical firms must spend resources on score inflation in order to protect their share of the prize. The leaderboard therefore distorts behaviour even when it does not create any asymmetry in outcomes.

Also note that equilibrium hacking exhibits: (i) *prize dependence*, firms hack more when prizes are larger; (ii) *ignorance of the future*, period τ hacking depends only on the size of current and past prizes $M_\tau, M_{\tau-1}, ..., M_1$; and (iii) *endowment effects*, a larger initial endowment reduces the amount of hacking in all future periods.

As a consequence of (ii), firms come to hack less over time for equal prize sizes: $h^*_{it} > h^*_{i,t+1}$ when $M_t = M_{t+1}$. Intuitively, firms use a non-zero proportion of the t period prize for research. As such, in the next period, they enter with a higher quality technology and less hacking is needed to reach the same leaderboard score. Relatedly, prizes are more valuable early on in the game. In the terminal period, one extra unit of budget is worth α because it is spent directly on research. If instead a dollar enters a contest period, it can again be turned into research, but now the resulting quality improvement also raises the firm's baseline leaderboard score by $\theta\alpha$. In equilibrium, that higher baseline means the firm needs $\theta\alpha/\gamma$ less hacking to reach the same score target, and those freed dollars are redirected to research. So one dollar entering a contest period is worth $\alpha + \alpha(\theta\alpha/\gamma) = \alpha/\gamma$. Each additional contest that lies ahead multiplies the value of a dollar by another factor of $1/\gamma$, which is why prizes won earlier compound over time.

4.1 Asymmetric Firms

Now suppose firms differ in research efficiency. Assume that firm 1 is a strictly more efficient firm than firm 2: $\alpha_1 > \alpha_2$ and equal initial budgets $B_{11} = B_{21} = B_1$. Since $\gamma_i \equiv 1 - \theta\alpha_i$, we have

$\gamma_1 < \gamma_2$. Theorem 1 therefore implies

$$\frac{s^*_{1t}}{s^*_{2t}} \;=\; \frac{B^*_{1,t+1}}{B^*_{2,t+1}} \;=\; \frac{\gamma_1}{\gamma_2} \;<\; 1 \qquad \text{for every contest date } t < T.$$

The inefficient firm posts the higher leaderboard score and receives the larger share of future funding in every interior contest period. In the first contest, where the firms begin from the same state, it also devotes a larger share of its budget to hacking:

$$h^*_{21} - h^*_{11} \;=\; \theta B_1 \left(\frac{\alpha_1}{\gamma_1} - \frac{\alpha_2}{\gamma_2} \right) \;>\; 0,$$

because $\alpha/(1-\theta\alpha)$ is strictly increasing in α. Intuitively, the inefficient firm has the lower marginal return to research, so it sacrifices less true quality when it shifts budget toward hacking. The efficient firm cannot match this manipulation without giving up a disproportionate amount of the activity in which it has comparative advantage. As a result, the inefficient firm inflates its measured score enough to overtake on the leaderboard.

Figure 2 gives the corresponding best-response geometry. Given that firm j's equilibrium strategy prescribes them to post a score s_j, firm i will respond with a leaderboard score of:

$$s^*_i(s_j) \;=\; \sqrt{M\gamma_i s_j} - s_j.$$

A higher γ_i shifts the best-response curve outward. Since $\gamma_2 > \gamma_1$, firm 2's best response lies above firm 1's, so firm 2 will choose to hack more aggressively and post a higher score for any given firm 1 strategy. Observe that, at the equilibrium point in figure 2, scores are strategic substitutes for firm 1. As a result, they reduce their hacking in response to firm 2's aggression.

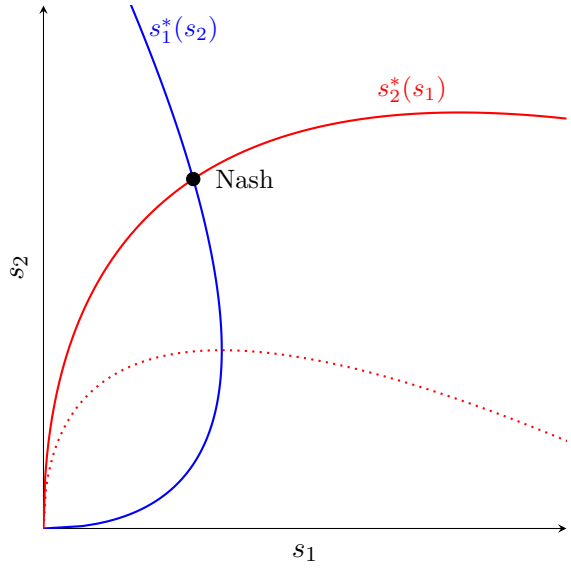


Figure 2: Best-response functions in a single leaderboard contest ($M = 1$). Firm 1's best response $s_1^*(s_2)$ (blue) corresponds to $\gamma_1 = 0.3$. The dotted red curve shows firm 2's best response in the symmetric case ($\gamma_2 = \gamma_1 = 0.3$); the solid red curve shows firm 2's best response when it has the stronger incentive to manipulate the leaderboard ($\gamma_2 = 0.7$). A higher γ shifts the best response outward, so the asymmetric equilibrium assigns firm 2 the higher score and the larger budget share.

This distortion has dynamic consequences. The higher leaderboard score delivers a larger share of the current prize, and that larger budget can then be used to sustain aggressive hacking in the next period. In this sense, leaderboards prevent productive firms from accessing the capital they need to innovate.

Finally, note that the wedge between truthful performance and leaderboard performance is larger when the leaderboard is more transparent. Since

$$\gamma_2 - \gamma_1 \;=\; \theta(\alpha_1 - \alpha_2),$$

a higher θ further increases firm 1's opportunity cost of hacking, relative to firm 2's. Interestingly, this implies that improving the design of the leaderboard to more accurately reflect true quality ($\uparrow \theta$), does not unambiguously improve things for the firms best at research.

Equilibrium payoffs. The next natural question to ask is how each firm performs. Under equal initial endowments, the equilibrium payoff of firm i can be written as

$$U_i^* = \underbrace{\frac{\alpha_i}{\gamma_i^{T-1}} B_1}_{\text{return on initial funding}} + \underbrace{\alpha_i \left[\frac{\gamma_i^2}{(\gamma_1+\gamma_2)^2} \sum_{t=1}^{T-1} \frac{M_t}{\gamma_i^{T-1-t}} \right]}_{\text{return on prizes}}. \tag{7}$$

The decomposition isolates two channels. The first term is the return on initial funding. It is always larger for the efficient firm: a higher α_i converts each unit of early budget into more true quality, and that early quality compounds through the remaining contests.

The second term is the return on prizes won along the equilibrium path. It is increasing in the prize sequence $\{M_t\}$, but its dependence on α_i is not monotone. A higher α_i raises the value of each prize dollar once won, but it also lowers the share of prizes the firm wins by reducing its leaderboard strength. Payoff reversal becomes possible when the strength effect causes the second term to be positive and dominate the first term.

Proposition 1. *Suppose $B_{11} = B_{21} = B_1$, $\alpha_1 > \alpha_2$, and the equilibrium path is interior. Then the inefficient firm earns the higher equilibrium payoff if and only if*

$$\sum_{t=1}^{T-1} \left(\alpha_2 \gamma_2^{t-T+3} - \alpha_1 \gamma_1^{t-T+3} \right) M_t > (\gamma_1+\gamma_2)^2 \left(\frac{\alpha_1}{\gamma_1^{T-1}} - \frac{\alpha_2}{\gamma_2^{T-1}} \right) B_1. \tag{8}$$

Appendix A.2 shows that this condition is just the exact sign restriction on $U_2^* - U_1^*$. It makes transparent that payoff reversal is a race between two forces: firm 1 earns more from any given unit of early funding, but firm 2 captures a larger share of prizes along the equilibrium path. The inequality also gives a useful interpretation. Payoff reversal is more likely when: (i) initial endowments are small relative to the prize pool, since the right-hand side scales linearly in B_1; (ii) more prize mass is placed on dates with larger coefficients $\alpha_2 \gamma_2^{t-T+3} - \alpha_1 \gamma_1^{t-T+3}$, which typically means later prizes because early dollars compound more strongly for the efficient firm; and (iii) the leaderboard distortion is strong enough that firm 2's systematic advantage in winning prizes outweighs firm 1's superior return from any given unit of funding.

Numerical Example. Consider a three period game $T = 3$. Let $\theta = 1$, $\alpha_1 = 0.7$, $\alpha_2 = 0.65$, $B_{11} = B_{21} = 1$, $M_1 = 10$, $M_2 = 60$, so $\gamma_1 = 0.3$ and $\gamma_2 = 0.35$. The interiority condition is satisfied

at both contest dates, and the inequality in Proposition 1 holds. In the first contest, firm 2 devotes 63% of its budget to hacking against firm 1's 15%, achieves a score $s^*_{21} = 0.870$ versus $s^*_{11} = 0.746$, and wins $B^*_{22} = 5.39$ against $B^*_{12} = 4.62$. The same pattern repeats in the second contest, and firm 2 enters the terminal period with a larger research budget, $B^*_{23} = 32.31$ against firm 1's $B^*_{13} = 27.69$. Terminal qualities are $U^*_1 \approx 21.70$ and $U^*_2 \approx 22.00$, so the inefficient firm ends the game with a better technology. Firm 1 holds the underlying quality lead all the way to the terminal period ($Q^*_{12} = 2.31$ against $Q^*_{22} = 1.00$); the ranking reverses only because the leaderboard shows firm 2 is winning and allocates it extra funding accordingly.

One might ask whether firm 1 can overturn the ranking simply by matching firm 2's aggressive hacking. Suppose firm 1 commits to $h_{1t} = h^*_{2t}$ in each contest period, and firm 2 best-responds. Because firm 1 has the higher α, devoting the same amount to hacking leaves it with higher underlying quality and therefore a higher leaderboard score, so firm 1 wins both contests and ends with $U_1^{\text{match}} \approx 21.54$ against firm 2's $U_2^{\text{match}} \approx 20.01$. But firm 1 is worse off than in equilibrium: its own terminal quality falls from 21.70 to 21.54. Matching sacrifices research precisely where firm 1 has its comparative advantage, so the aggressive strategy dominates the opponent only by burning more of firm 1's surplus than firm 2's. In equilibrium, the efficient firm rationally cedes the leaderboard rather than fighting the contest on the inefficient firm's terms.

4.2 Long-run Leaderboard Dynamics

The short-run analysis shows why an inefficient firm can top the leaderboard even though it is worse at research. The long-run question is whether the efficient firm eventually catches and overtakes it on the leaderboard.

To study that question it is useful to focus on

$$X_{it} \equiv Q_{i,t-1} + \alpha_i B_{it},$$

X_{it} represents the quality that firm i has already built plus the quality it could add immediately if it devoted its current budget entirely to research. This is the relevant state for leaderboard dynamics because θX_{it} is the firm's truthful leaderboard score in period t. If this truthful component rises fast enough, the efficient firm's organic score eventually catches up and exceeds the inefficient firm's

maximum score under full hacking. If it does not, the inefficient firm can keep buying enough score to remain on top.

It is also helpful to parametrize prize growth explicitly. Let

$$M_t = M_1 \lambda^{t-1},$$

where λ is the gross growth rate of the prize sequence: $\lambda = 1$ means constant prizes, while $\lambda > 1$ means that the pool of capital allocated through the leaderboard expands over time. Appendix A.3 shows that along any interior path,

$$X_{i,t+1} \;=\; \frac{X_{it}}{\gamma_i} \;+\; \frac{\alpha_i \gamma_i^2}{(\gamma_1 + \gamma_2)^2}\, M_t.$$

So the long-run issue is whether underlying capability grows fast enough relative to the prize sequence to make manipulation unattractive. Equivalently, dividing by M_t gives a linear recursion for the ratio X_{it}/M_t: that ratio diverges when $\lambda \leq 1/\gamma_i$ and converges when $\lambda > 1/\gamma_i$. This normalization compares the firm's cumulative truthful state to the scale of the contest it is trying to win. Slow prize growth lets genuine quality accumulate faster than the value of score inflation, whereas fast prize growth can keep manipulation valuable indefinitely. The following propositions study the sequence of T-period games as the horizon $T \to \infty$.

Constant prizes. Constant prizes are the clean self-correction case. When $\lambda = 1$, the return to winning the next contest stays tied to the same prize M in every period. By contrast, research raises true quality permanently, so its effect carries forward into all later contests. Early in the race, firm 2 may still find it worthwhile to sacrifice research for hacking in order to win funding. But over time the cumulative truthful component of the leaderboard becomes larger, while the gain from manipulating a single contest does not scale up with it. That is why deception can be profitable at first but cannot remain the dominant strategy forever.

Proposition 2. *Under constant prizes the inefficient firm cannot remain at the top of the leaderboard forever. Assume $M_t \equiv M > 0$, $B_{11} = B_{21} = B_1$, and $\alpha_1 > \alpha_2$. Then there exist finite dates τ_H and $\tau_S \geq \tau_H$ such that $h^*_{1t} = h^*_{2t} = 0$ for all $t \geq \tau_H$ and $\underline{s}_{1t} > \overline{s}_{2t}$ for all $t \geq \tau_S$.*

The logic is that constant prizes keep the per-period value of leaderboard manipulation from growing over time, while research has a cumulative effect on future truthful performance. Early on, firm 2 can profitably inflate its score to win funding. But as both firms accumulate genuine quality, the truthful component of the leaderboard becomes large relative to the fixed prize at stake in each contest. At that point, diverting budget into hacking is no longer worth it. Appendix A.4.2 formalises this first step by showing that there is a finite date after which full research is the unique best response for both firms, so the race enters an honest tail.

From that point onward the leaderboard simply tracks true quality, and future funding is allocated on the basis of genuine capability. That makes the ratio Q_{1t}/Q_{2t} the natural object to follow. Appendix A.4.3 shows that this ratio rises monotonically once hacking has disappeared, because every pound allocated to firm 1 produces more future quality than the same pound allocated to firm 2. So even if firm 2 is still ahead when the honest tail begins, firm 1 keeps closing the gap and eventually overtakes on the truthful leaderboard.

Once firm 1 has overtaken, the dynamic turns decisively in its favour. A leader on the truthful leaderboard receives at least half of each subsequent prize, and each of those pounds is more productive in firm 1's hands because $\alpha_1 > \alpha_2$. The true-quality gap therefore keeps widening, until after finitely many further contests there is clear air on the leaderboard:

$$\underline{s}_{1t} > \overline{s}_{2t} \qquad \forall\, t \geq \tau_S$$

Figure 3 illustrates this final step: after the honest tail begins, firm 1's truthful lower bound keeps rising while firm 2's maximal feasible score rises more slowly, and the two cross at the separation date τ_S. At that point the inefficient firm cannot reclaim the top spot even with maximal hacking.

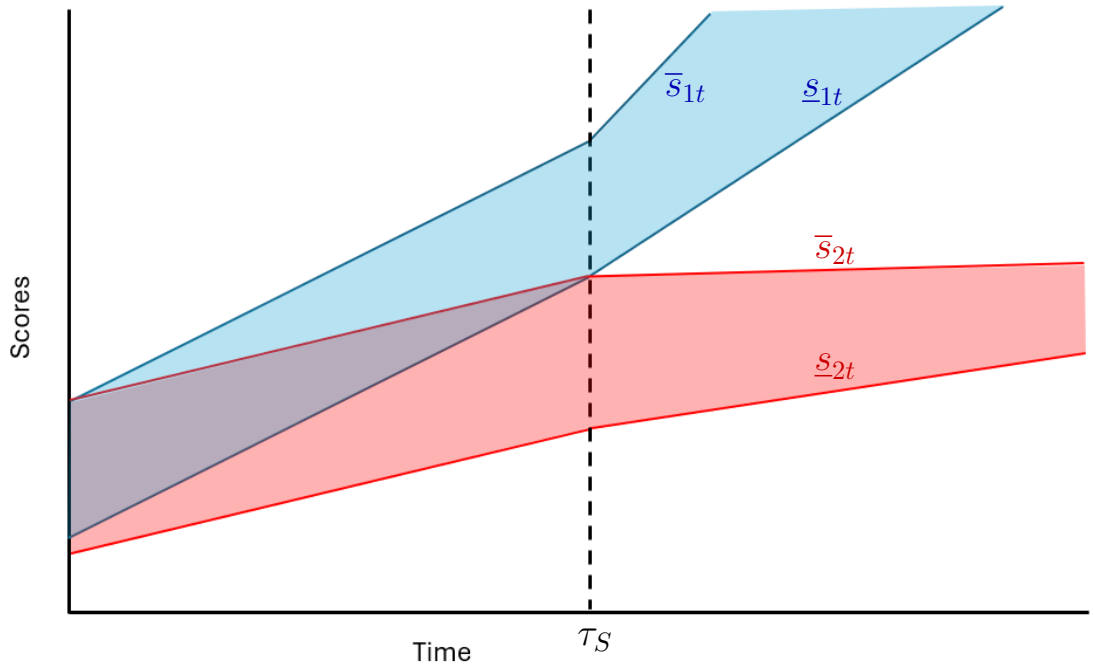


Figure 3: Stylized score dynamics under constant prizes. The blue region shows firm 1's feasible score interval $[\underline{s}_{1t}, \overline{s}_{1t}]$ and the red region shows firm 2's feasible score interval $[\underline{s}_{2t}, \overline{s}_{2t}]$. At the date τ_S, firm 1's truthful lower bound crosses above firm 2's maximal feasible score, so to the right of τ_S the two feasible regions are fully separated. After separation, firm 1 begins to win an increasing share of the prize and thus their budget (and feasible interval) widens, while firm 2's narrows.

Appendix A.4.5 also gives a simple bound on the waiting time until separation. If τ_O is the first contest date in the honest tail at which firm 1 weakly overtakes on the truthful leaderboard, then $\tau_S \leq \tau_O + 1 + n_S$, where n_S is the smallest number of additional contests required for firm 1's true-quality lead to exceed the separation threshold. Formally, if $D_t \equiv Q_{1t} - Q_{2t}$, then $n_S \equiv \min\left\{n \geq 0 : D_{\tau_O} + n\frac{M}{2}(\alpha_1 - \alpha_2) > \frac{\gamma_1 M}{2\theta}\right\}$.

Separation is faster when the efficiency gap $\alpha_1 - \alpha_2$ is larger, when firm 1 is already close to the lead when the honest tail begins, and when the leaderboard is more transparent.

Growing prizes. The opposite force appears when the prize sequence itself grows over time. If $M_t = M_1\lambda^{t-1}$ with $\lambda > 1$, then the amount of capital allocated through the leaderboard expands every period. That helps the efficient firm build genuine capability, but it also raises the score

that must be posted to stay at the top of the leaderboard. Prize growth can therefore prevent self-correction because it raises both the return to winning the leaderboard and the resources available to the current leader, allowing the inefficient firm to finance persistent score inflation even as the efficient firm's true quality improves.

Proposition 3 (Perpetual leaderboard deception under fast prize growth)**.** *Assume* $M_t = M_1\lambda^{t-1}$, $B_{11} = B_{21} = B_1$, $\alpha_1 > \alpha_2$, *and that the first contest date is interior. If*

$$\frac{1}{\gamma_1} + \frac{\theta\alpha_1}{\gamma_2} \;<\; \lambda$$

then the equilibrium path remains interior at every contest date. Consequently, $s^*_{2t} > s^*_{1t}$ *and* $B^*_{2,t+1} > B^*_{1,t+1}$ *at every contest date* t.

This bound ensures that firm 1's truthful leaderboard score never catches the interior target. Here, the same interior logic that gives firm 2 the short-run advantage continues to operate forever: the efficient firm keeps building more capability, but not quickly enough to overtake on the leaderboard, while the inefficient firm keeps converting its lower opportunity cost of hacking into a persistent lead. In that region the honest tail never arrives, because the leaderboard keeps replenishing the resources needed to sustain manipulation.

Remark 1 (Explosive prize growth). Technically, there is also an upper bound on the growth rate: $\lambda < \frac{1}{\gamma_1} + \frac{1}{\gamma_2}$. If prize growth exceeds this, both firms find it optimal to hack only. The leaderboard then becomes closer to a spending contest whilst quality plateaus. I do not pursue that case here.

5 Welfare and Leaderboard Design

This section asks whether the presence of the leaderboard improves or worsens total research. The answer depends on how quickly the prize pool grows. With flat prizes, the leaderboard is initially wasteful but eventually useful once true capability catches up and the ranking becomes effectively honest. With sufficiently fast prize growth, that self-correction never arrives.

Two Allocation Mechanisms

We compare the leaderboard to a natural counterfactual.

Leaderboard (LB). Prizes are allocated via the Tullock contest over leaderboard scores, as in Section 3. Along an interior path, firm i receives share $\sigma_{it}^{\mathrm{LB}} = \gamma_i/(\gamma_1+\gamma_2)$ and both firms hack strictly.

Random (RND). Each prize M_t is split equally: $B_{i,t+1} = M_t/2$. Hacking yields no return, so $h_{it}^* = 0$ and all budget flows into research.

Let $W \equiv Q_{1T} + Q_{2T}$ denote aggregate terminal capability and be our welfare criterion. Writing $\sigma_{it} \equiv B_{i,t+1}/M_t$ for firm i's share of prize M_t and $\Delta_t \equiv \alpha_1 h_{1t}^* + \alpha_2 h_{2t}^*$ for the capability lost to hacking in contest t, aggregate welfare can be written as

$$W \;=\; (\alpha_1+\alpha_2)B_1 \;+\; \sum_{t=1}^{T-1} M_t\left(\alpha_1\sigma_{1t} + \alpha_2\sigma_{2t}\right) \;-\; \sum_{t=1}^{T-1}\Delta_t. \tag{9}$$

Relative to RND, contest t contributes

$$\delta_t \;\equiv\; M_t(\alpha_1-\alpha_2)\big(\sigma_{1t}^{\mathrm{LB}} - \tfrac{1}{2}\big) \;-\; \Delta_t^{\mathrm{LB}}, \qquad W^{\mathrm{LB}}(T) - W^{\mathrm{RND}}(T) \;=\; \sum_{t=1}^{T-1}\delta_t.$$

This decomposition isolates the two margins in the model. The first term $M_t(\alpha_1-\alpha_2)\big(\sigma_{1t}^{\mathrm{LB}} - \frac{1}{2}\big)$ is a misallocation effect: the leaderboard helps total innovation only when the efficient firm receives more than half of the prize. The second term Δ_t^{LB} is pure waste from hacking and only appears where there is a leaderboard to manipulate. In the interior regime both forces are negative. Once the leaderboard becomes honest, hacking disappears and only the allocation term remains. From that point onward welfare is weakly negative until firm 1 overtakes on the truthful leaderboard, and strictly positive thereafter.

Welfare Frontier

Proposition 4 (Welfare under flat and fast-growing prize sequences)**.** *Assume $\alpha_1 > \alpha_2$, $B_{11} = B_{21} = B_1$, and that the first contest date is interior.*

(i) *If prizes are flat, $M_t \equiv M > 0$, then the leaderboard is welfare-inferior at short horizons but welfare-superior at long horizons: there exist $T^- < T^+$ such that $W^{\mathrm{LB}}(T) < W^{\mathrm{RND}}(T)$ for all $T \le T^-$ and $W^{\mathrm{LB}}(T) > W^{\mathrm{RND}}(T)$ for all $T \ge T^+$.*

(ii) *If prizes grow geometrically,* $M_t = M_1\lambda^{t-1}$*, and growth lies in the perpetual-deception region*

$$\frac{1}{\gamma_1} + \frac{\theta\alpha_1}{\gamma_2} \;<\; \lambda$$

then $W^{\mathrm{LB}}(T) < W^{\mathrm{RND}}(T)$ *for every* $T \geq 2$*. Moreover,* $W^{\mathrm{RND}}(T) - W^{\mathrm{LB}}(T) \to \infty$ *as* $T \to \infty$*.*

Proof sketch. Appendix A.7 combines the accounting identity above with the long-run results from Section 4. Under flat prizes, Proposition 2 implies that the manipulable phase ends in finite time and that firm 1 eventually overtakes on the truthful leaderboard. Before that date each δ_t is weakly negative; afterwards each is strictly positive. The early loss is therefore finite, whereas the positive tail accumulates one constant-prize gain after another, so the sign eventually flips.

Under fast prize growth, Proposition 3 implies that the interior regime persists forever. Then every period carries the same two welfare distortions as in the short run: firm 1 receives less than half of the prize and both firms hack. So every δ_t is strictly negative. Because the misallocation term is proportional to M_t and the prize sequence grows geometrically, the cumulative welfare shortfall diverges.

Optimal prize timing. Within the interior region, the welfare identity (9) reveals that the misallocation channel depends only on the total prize pool $\sum M_t$, because $\sigma_{it}^{\mathrm{LB}} = \gamma_i/(\gamma_1+\gamma_2)$ is constant across contest dates. Prize timing therefore affects welfare only through the hacking waste channel.

The key is the cross-period dampening identified in Section 4: a larger prize in period t funds more genuine research in period $t+1$, which raises the baseline score $\theta Q_{i,\tau-1}$ entering every later contest $\tau > t$. A higher baseline reduces the hacking required to reach the equilibrium target score, so Δ_τ falls. A dollar allocated to M_t therefore reduces waste in all $T-1-t$ subsequent periods; a dollar in $M_{t'}$ for $t' > t$ reduces waste only in the shorter tail of $T-1-t'$ remaining periods. Shifting prize money earlier therefore reduces total waste without changing misallocation.

Proposition 5 (Front-loading is optimal)**.** *Fix the total prize budget* $\bar{M} = \sum_{t=1}^{T-1} M_t$*, and restrict attention to prize sequences for which the equilibrium path is interior at every contest date. Let* M *be such a sequence. If* $t < t'$ *and* $\varepsilon > 0$ *is small enough that the perturbed sequence* $\tilde{M}$ *defined*

by $\tilde{M}_t = M_t + \varepsilon$, $\tilde{M}_{t'} = M_{t'} - \varepsilon$, *and* $\tilde{M}_\tau = M_\tau$ *for* $\tau \notin \{t, t'\}$ *also remains interior, then* $W^{\mathrm{LB}}(\tilde{M}) > W^{\mathrm{LB}}(M)$. *Consequently, every welfare-maximising interior prize sequence is weakly front-loaded:* $M_1^* \geq M_2^* \geq \cdots \geq M_{T-1}^*$.

The proof, in Appendix A.6, differentiates the welfare identity using the closed-form dependence of Δ_τ on the entire prize sequence. Holding $\bar{M}$ fixed, shifting prize mass from a later date to an earlier one leaves the interior misallocation term unchanged and strictly lowers cumulative hacking waste.

Remark 2 (Planner's dilemma). Front-loading reduces waste but can worsen the distributional distortion. From the payoff formula (6), shifting a dollar from M_{t+1} to M_t changes firm i's payoff by $\theta\alpha_i^2\gamma_i^{t+3-T}/(\gamma_1+\gamma_2)^2$, so the differential benefit accruing to firm 2 rather than firm 1 is

$$\frac{\theta\left[\alpha_2^2\gamma_2^{t+3-T} - \alpha_1^2\gamma_1^{t+3-T}\right]}{(\gamma_1+\gamma_2)^2}.$$

For the latest reallocation ($t = T-2$), this expression is positive if and only if $\alpha_2^2(1-\theta\alpha_2) > \alpha_1^2(1-\theta\alpha_1)$. A sufficient, but not necessary, condition is $\alpha_2 > 2/(3\theta)$, since then both firms lie on the decreasing branch of $\alpha^2(1-\theta\alpha)$. In that case the waste-minimising reallocation of the last two prizes also widens the payoff gap in favour of the inefficient firm, making payoff reversal more likely within the interior region. A designer who simultaneously wants low hacking waste *and* a correctly ranked winner therefore faces a genuine tension. The tension eases for earlier reallocations, where the compounding factor γ_i^{t+3-T} dominates and the efficient firm becomes the larger beneficiary.

6 Sequential Leaderboard Publication

In many leaderboard-driven industries, scores are not submitted simultaneously. AI labs often release results one by one and public leaderboards update immediately, so later movers can condition their submission on what is already on the board and early movers have commitment power. Does this timing relieve or exacerbate the distortion?

Helpfully, the game remains tractable because the continuation value remains affine in inherited quality and current budget, so the same backward-induction argument applies: each contest subgame reduces to a static Stackelberg leaderboard contest with leader cost $1/\gamma_L$ and follower

cost $1/\gamma_F$. The follower best-responds to the leader's posted score, and the leader chooses its score anticipating that response. The full characterization, including the interior Stackelberg score and budget formulas and the follower-participation condition $\gamma_L < 2\gamma_F$, is given in Appendix A.8. The rest of this section reads off the economic implications.

Symmetric firms

When $\alpha_1 = \alpha_2$ and initial endowments are equal, leader and follower post identical scores, win identical budget shares, and choose identical research and hacking in every contest period. Timing alone creates neither a score gap nor a budget gap, and the sequential leaderboard coincides exactly with the simultaneous leaderboard of Section 4. Sequential publication only matters when firms are asymmetric.

Notation. For the asymmetric case $\alpha_1 > \alpha_2$ (so $\gamma_1 < \gamma_2$), let superscript E denote the Stackelberg regime in which the *efficient* firm (firm 1) leads, and I the regime in which the *inefficient* firm (firm 2) leads. Write $\beta_i^E \equiv B_{i,t+1}^E / M_t$ for firm i's budget share and $H^E \equiv \alpha_1 h_{1t}^E + \alpha_2 h_{2t}^E$ for aggregate hacking waste in the E-regime, and likewise for I. The simultaneous-play leaderboard carries no superscript.

Efficient Firm Leads

Two effects appear, and they work against each other.

The score gap widens. Firm 2's lead on the leaderboard becomes more pronounced than under simultaneous play:

$$\frac{s_{2t}^E}{s_{1t}^E} = \frac{2\gamma_2 - \gamma_1}{\gamma_1} > \frac{\gamma_2}{\gamma_1} = \frac{s_{2t}}{s_{1t}}.$$

Because the Tullock allocation maps score ratios directly into budget ratios, the funding gap widens correspondingly ($\beta_2^E > \beta_2$). The mechanism: the efficient firm has a high marginal cost of posting score ($1/\gamma_1 > 1/\gamma_2$), so when it commits first it commits to a low target; the follower best-responds with a low score too, but still high enough to walk away with an even larger share of the prize than under simultaneous play. First-mover commitment is used by the efficient firm to concede the leaderboard rather than to seize it.

Total hacking falls. At a fixed continuation state, $H^E < H$. The leader hacks $(\gamma_2 - \gamma_1)M_t/(4\gamma_2)$ less than it would against a clone with the same γ_1, and the follower hacks $(\gamma_2 - \gamma_1)^2 M_t/(4\gamma_2^2)$ less than against a γ_2-clone. The leader's low score draws a low follower response, so the overall scale of the contest shrinks.

Efficient-first sequentiality therefore trades off the two welfare channels of Section 5 against each other: the hacking-waste channel improves ($H^E < H$) while the misallocation channel worsens ($\beta_2^E > \beta_2$). The net effect on aggregate capability depends on the size of the prize pool relative to the efficiency gap.

Inefficient Firm Leads

If instead firm 2 moves first and the follower-participation condition $\gamma_2 < 2\gamma_1$ holds, both channels move the wrong way.

The score gap widens further. Firm 2's leaderboard lead is now the largest of the three timing regimes:

$$\frac{s_{2t}^I}{s_{1t}^I} = \frac{\gamma_2}{2\gamma_1 - \gamma_2} > \frac{2\gamma_2 - \gamma_1}{\gamma_1} > \frac{\gamma_2}{\gamma_1}.$$

The funding gap widens correspondingly, with $\beta_2^I > \beta_2^E > \beta_2$. Now it is the *inefficient* firm that commits to a score first, and because its low marginal cost of posting score ($1/\gamma_2 < 1/\gamma_1$) makes the commitment cheap, it uses that commitment to seize an even bigger slice than in simultaneous play.

Total hacking rises. At a fixed continuation state, $H^I > H$. The inefficient leader hacks $(\gamma_2 - \gamma_1)M_t/(4\gamma_1)$ *more* than against a γ_2-clone; the efficient follower hacks $(\gamma_2 - \gamma_1)^2 M_t/(4\gamma_1^2)$ less. The leader's incremental aggression outweighs the follower's retreat, so total manipulation expands.

Inefficient-first sequentiality is therefore the worst of the three timing arrangements: both the misallocation channel and the hacking-waste channel deteriorate relative to simultaneous play, with no offsetting force.

Takeaway. Sequential leaderboards are neutral when firms are symmetric, because the dynamic recursion reproduces the same outcome as the simultaneous leaderboard. With heterogeneous firms,

however, timing never improves selection: the inefficient firm still receives too much funding, and the distortion is larger when it moves first. What does depend on the mover order is hacking. If the efficient firm leads, sequential publication lowers total manipulation but worsens allocation. If the inefficient firm leads, sequential publication worsens both margins. So sequential leaderboards are most harmful precisely when they allow the natural leaderboard hacker to post first.

7 Rational Demand for a Gameable Leaderboard

The main model takes as primitive that leaderboard-score shares allocate next-period capital. An obvious objection is that: if inefficient firms are incentivised to post higher scores, why should consumers reward high scores with their demand? This section gives a short demand-side foundation for the reduced form Tullock contest in Section 3, building on the frameworks established in Jia [2008] and Frankel and Kartik [2019]. The key insight is that, a high score can be achieved by genuinely good firms for free, and bad firms at a cost. Consequently, some types are able to imitate high scores but others are separated simultaneously, resulting in scores representing good news *on average* and a partially informative leaderboard.

Consider a one-period auxiliary market with many symmetric firms. Firm i has privately observed type (Q_i, g_i), where $Q_i > 0$ is true quality and $g_i > 0$ measures the productivity of hacking. It chooses hacking $h_i \geq 0$, posts the public score

$$s_i = Q_i + h_i,$$

and pays cost $h_i^2/(2g_i)$. Let (Q_i, g_i) have means (μ_Q, μ_g), variances (σ_Q^2, σ_g^2), covariance σ_{Qg}, and a nonsingular covariance matrix. Consumers observe scores, not types, and form the posterior mean

$$\widehat{Q}(s_i) \equiv \mathbb{E}[Q_i \mid s_i].$$

By Jia [2008], random-utility demand with Fréchet shocks delivers the ratio contest-success form. Thus if perceived qualities are $\widehat{Q}(s_i) > 0$, aggregate demand for firm i can be written as

$$D_i(s) = M \frac{\widehat{Q}(s_i)}{\sum_k \widehat{Q}(s_k)}. \tag{10}$$

When the posterior is affine in the score, perceived quality is itself an affine transformation of the public leaderboard score. Normalising the public score in perceived-quality units therefore gives the Tullock allocation rule used in Section 3.

It remains to show that an affine posterior can have a strictly positive slope even though firms endogenously hack. I use the linear-posterior result of Frankel and Kartik [2019]: under the elliptical specification, the posterior mean in a linear equilibrium is affine,

$$\widehat{Q}(s) = \mu_Q + \lambda(s - \mu_s),$$

where consistency requires $\lambda = \mathrm{Cov}(Q, s)/\mathrm{Var}(s)$.

Proposition 6 (Positive leaderboard responsiveness)**.** *Suppose $\sigma_Q^2 > 0$, $\sigma_g^2 > 0$, and $\sigma_Q^2\sigma_g^2 - \sigma_{Qg}^2 > 0$. In the affine-posterior class described above, there exists an equilibrium with $\lambda > 0$. In any such equilibrium, the score has partial but imperfect informational content:*

$$R^2 \equiv \mathrm{Corr}(Q_i, s_i)^2 \in (0, 1).$$

Proof. In a large symmetric market, each firm takes the denominator in (10) as fixed. Since $\mathbb{E}[\widehat{Q}(s_i)] = \mu_Q$, write the marginal value of perceived quality as $\eta \equiv M/(N\mu_Q) > 0$. Given an affine posterior with slope λ, firm i solves

$$\max_{h_i \geq 0} \ \eta\widehat{Q}(Q_i + h_i) - \frac{h_i^2}{2g_i}.$$

This problem is strictly concave in h_i, and when $\lambda > 0$ its interior solution is

$$h_i^*(\lambda) = \eta\lambda g_i.$$

The induced score is therefore $s_i = Q_i + \eta\lambda g_i$. Posterior consistency requires

$$\lambda = \frac{\mathrm{Cov}(Q_i, s_i)}{\mathrm{Var}(s_i)} = \frac{\sigma_Q^2 + \eta\lambda\sigma_{Qg}}{\sigma_Q^2 + 2\eta\lambda\sigma_{Qg} + \eta^2\lambda^2\sigma_g^2}. \tag{11}$$

Equivalently, λ must solve

$$F(\lambda) \equiv \eta^2\sigma_g^2\lambda^3 + 2\eta\sigma_{Qg}\lambda^2 + (\sigma_Q^2 - \eta\sigma_{Qg})\lambda - \sigma_Q^2 = 0.$$

Now $F(0) = -\sigma_Q^2 < 0$, while $F(\lambda) \to \infty$ as $\lambda \to \infty$ because $\eta^2\sigma_g^2 > 0$. By continuity, there is at least one positive solution $\lambda^* > 0$. At such a solution, firms choose $h_i^* = \eta\lambda^* g_i > 0$ and consumers' affine posterior slope is consistent with the induced score distribution, giving a positive-λ equilibrium.

Finally, at any positive solution to (11),

$$\mathrm{Cov}(Q_i, s_i) = \lambda \,\mathrm{Var}(s_i) > 0,$$

so $R^2 > 0$. Cauchy's inequality gives $R^2 \leq 1$. The covariance matrix of (Q_i, g_i) is nonsingular and $\eta\lambda > 0$, so $s_i = Q_i + \eta\lambda g_i$ is not almost surely an affine transformation of Q_i. The inequality is therefore strict: $R^2 < 1$. □

The proposition is the demand-side analogue of the model's main trade-off. The leaderboard is informative enough that rational consumers respond to it, but the response induces hacking, so the score remains an imperfect signal of true quality.

8 Conclusion

This paper studies an R&D race in which a public leaderboard measures performance and allocates the resources that finance future innovation. When leaderboard scores can be manipulated, the leaderboard creates a multitasking problem between durable research and transitory score inflation. This distortion is most severe for the firms that are best at genuine research, because hacking is most costly exactly where the return to research is highest. As a result, the leaderboard can direct funding toward less efficient firms, and in sufficiently fast-growing environments it can sustain their lead for a long time.

The welfare consequences therefore depend on whether genuine technological improvements can outpace the influx of capital that sustains rent-seeking score inflation. With flat prizes, the early phase of manipulation is eventually exhausted and the leaderboard becomes useful once the more

productive firm overtakes on the truthful margin. With sufficiently fast prize growth, that correction never arrives and the leaderboard remains welfare-reducing. The evidence from the Hugging Face reranking is consistent with the allocative premise behind the model: changes in rank appear to reallocate later adoption even when the underlying models are unchanged. Taken together, these results suggest that leaderboard design should be evaluated not only by how informative a ranking is at a point in time, but also by how easily it can be manipulated and how strongly it feeds back into future resources.

References

Norah Alzahrani, Hisham Abdullah Alyahya, Yazeed Alnumay, Sultan Alrashed, Shaykhah Alsubaie, Yusef Almushaykeh, Faisal Mirza, Nouf Alotaibi, Nora Altwairesh, Areeb Alowisheq, M. Saiful Bari, and Haidar Khan. When Benchmarks are Targets: Revealing the Sensitivity of Large Language Model Leaderboards, July 2024. URL http://arxiv.org/abs/2402.01781. arXiv:2402.01781 [cs].

Masaki Aoyagi. Information feedback in a dynamic tournament. *Games and Economic Behavior*, 70(2):242–260, November 2010. ISSN 0899-8256. doi: 10.1016/j.geb.2010.01.013. URL https://www.sciencedirect.com/science/article/pii/S0899825610000357.

George Baker, Robert Gibbons, and Kevin J. Murphy. Subjective Performance Measures in Optimal Incentive Contracts. *The Quarterly Journal of Economics*, 109(4):1125–1156, 1994. ISSN 0033-5533. doi: 10.2307/2118358. URL https://www.jstor.org/stable/2118358.

Jean-Michel Benkert and Igor Letina. Designing Dynamic Research Contests. *American Economic Journal: Microeconomics*, 12(4):270–289, November 2020. ISSN 1945-7669. doi: 10.1257/mic.20180263. URL https://www.aeaweb.org/articles?id=10.1257/mic.20180263.

Yatong Chen, Guanhua Zhang, and Moritz Hardt. Leaderboard Incentives: Model Rankings under Strategic Post-Training, March 2026. URL http://arxiv.org/abs/2603.08371. arXiv:2603.08371 [cs].

Luis C. Corchón. The theory of contests: a survey. *Review of Economic Design*, 11(2):69–100, September 2007. ISSN 1434-4742, 1434-4750. doi: 10.1007/s10058-007-0032-5. URL http://link.springer.com/10.1007/s10058-007-0032-5.

David Dranove, Daniel Kessler, Mark McClellan, and Mark Satterthwaite. Is More Information Better? The Effects of "Report Cards" on Health Care Providers. *Journal of Political Economy*, 111(3):555–588, 2003. ISSN 0022-3808. doi: 10.1086/374180. URL https://www.jstor.org/stable/10.1086/374180.

Alexander Erlei, Federico Cau, Radoslav Georgiev, Sagar Kumar, Kilian Bizer, and Ujwal Gadiraju. When Life Gives You AI, Will You Turn It Into A Market for Lemons? Understanding How

Information Asymmetries About AI System Capabilities Affect Market Outcomes and Adoption, January 2026. URL http://arxiv.org/abs/2601.21650. arXiv:2601.21650 [cs].

Alex Frankel and Navin Kartik. Muddled Information. *Journal of Political Economy*, 127(4):1739–1776, August 2019. doi: 10.1086/701604. URL https://www.journals.uchicago.edu/doi/abs/10.1086/701604.

Marina Halac, Navin Kartik, and Qingmin Liu. Contests for Experimentation. *Journal of Political Economy*, 125(5):1523–1569, 2017. ISSN 0022-3808. URL https://www.jstor.org/stable/26550449.

Christopher Harris and John Vickers. Racing with Uncertainty. *The Review of Economic Studies*, 54 (1):1–21, 1987. ISSN 0034-6527. doi: 10.2307/2297442. URL https://www.jstor.org/stable/2297442.

Toomas Hinnosaar. Optimal sequential contests. *Theoretical Economics*, 19(1):207–244, 2024. ISSN 1933-6837. doi: 10.3982/TE5536. URL https://econtheory.org/ojs/index.php/te/article/view/5536.

Bengt Holmstrom and Paul Milgrom. Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design. *Journal of Law, Economics, & Organization*, 7:24–52, 1991. URL http://www.jstor.org/stable/764957.

Stanton Hudja, Brian Roberson, and Yaroslav Rosokha. Public Leaderboard Feedback in Sampling Competition: An Experimental Investigation. *The Review of Economics and Statistics*, 107(2): 555–569, 2025. URL https://ideas.repec.org//a/tpr/restat/v107y2025i2p555-569.html.

Hao Jia. A Stochastic Derivation of the Ratio Form of Contest Success Functions. *Public Choice*, 135(3/4):125–130, 2008. ISSN 0048-5829. URL https://www.jstor.org/stable/27698259.

Jorge Lemus and Guillermo Marshall. Dynamic Tournament Design: Evidence from Prediction Contests. *Journal of Political Economy*, 129(2):383–420, February 2021. ISSN 0022-3808, 1537-534X. doi: 10.1086/711762. URL https://www.journals.uchicago.edu/doi/10.1086/711762.

Adam Meirowitz. Electoral Contests, Incumbency Advantages, and Campaign Finance. *The Journal of Politics*, 70(3):681–699, July 2008. ISSN 0022-3816. doi: 10.1017/S0022381608080699. URL https://www.journals.uchicago.edu/doi/abs/10.1017/S0022381608080699.

Simon Ott, Adriano Barbosa-Silva, Kathrin Blagec, Jan Brauner, and Matthias Samwald. Mapping global dynamics of benchmark creation and saturation in artificial intelligence. *Nature Communications*, 13:6793, November 2022. ISSN 2041-1723. doi: 10.1038/s41467-022-34591-0. URL https://pmc.ncbi.nlm.nih.gov/articles/PMC9649641/.

Rolf Pfister and Hansueli Jud. Understanding and Benchmarking Artificial Intelligence: OpenAI's o3 Is Not AGI, January 2025. URL http://arxiv.org/abs/2501.07458. arXiv:2501.07458 [cs].

Inioluwa Deborah Raji, Emily M. Bender, Amandalynne Paullada, Emily Denton, and Alex Hanna. AI and the Everything in the Whole Wide World Benchmark, November 2021. URL http://arxiv.org/abs/2111.15366. arXiv:2111.15366 [cs].

Jennifer F. Reinganum. A Dynamic Game of R and D: Patent Protection and Competitive Behavior. *Econometrica*, 50(3):671–688, 1982. ISSN 0012-9682. doi: 10.2307/1912607. URL https://www.jstor.org/stable/1912607.

Shivalika Singh, Yiyang Nan, Alex Wang, Daniel D'Souza, Sayash Kapoor, Ahmet Üstün, Sanmi Koyejo, Yuntian Deng, Shayne Longpre, Noah A. Smith, Beyza Ermis, Marzieh Fadaee, and Sara Hooker. The Leaderboard Illusion, May 2025. URL http://arxiv.org/abs/2504.20879. arXiv:2504.20879 [cs].

Huseyin Yildirim. Contests with multiple rounds. *Games and Economic Behavior*, 51(1):213–227, 2005. URL https://www.sciencedirect.com/science/article/pii/S0899825604000673.

A Proofs

A.1 Proof of Theorem 1

Fix a contest period t and a continuation subgame in which the interiority condition holds. The proof proceeds by backward induction. We show that the continuation payoff of firm i, viewed as a function of the state (Q, B) entering any period t, takes the affine form

$$V_i^t(Q, B) \;=\; \frac{Q + \alpha_i B}{\gamma_i^{T-t}} + D_i^t, \tag{12}$$

where D_i^t depends on the remaining prizes $\{M_\tau\}_{\tau \geq t}$ but not on the state. Here Q denotes accumulated quality entering period t (that is, $Q = Q_{i,t-1}$) and B denotes the budget available for allocation in period t (that is, $B = B_{it}$).

Base case ($t = T$). In the terminal period there is no contest, so firm i invests its entire budget in research: $r_{iT} = B_{iT}$. Terminal quality is $Q_{iT} = Q_{i,T-1} + \alpha_i B_{iT}$, yielding

$$V_i^T(Q, B) \;=\; Q + \alpha_i B.$$

This is (12) with $T - t = 0$ and $D_i^T = 0$.

Inductive step. Suppose (12) holds at $t + 1$. Firm i enters period $t \in \{1, \ldots, T-1\}$ with state $(Q_{i,t-1}, B_{it})$ and chooses (r_{it}, h_{it}) subject to $r_{it} + h_{it} = B_{it}$ and $r_{it}, h_{it} \geq 0$. These choices determine

$$\begin{aligned} Q_{it} &= Q_{i,t-1} + \alpha_i r_{it}, \\ s_{it} = \theta Q_{it} + h_{it} \;&=\; \theta Q_{i,t-1} + \theta \alpha_i r_{it} + (B_{it} - r_{it}) \;=\; \theta Q_{i,t-1} + B_{it} - \gamma_i r_{it}, \\ B_{i,t+1} &= M_t \, \frac{s_{it}}{s_{it} + s_{jt}}. \end{aligned}$$

Step 1: Reducing the objective. By the inductive hypothesis,

$$V_i^{t+1}(Q_{it}, B_{i,t+1}) \;=\; \frac{Q_{it} + \alpha_i B_{i,t+1}}{\gamma_i^{T-t-1}} + D_i^{t+1}.$$

Since $\gamma_i^{-(T-t-1)} > 0$ and D_i^{t+1} is independent of firm i's period-t action, maximising V_i^{t+1} is equivalent to maximising

$$Q_{it} + \alpha_i B_{i,t+1} \;=\; Q_{i,t-1} + \alpha_i r_{it} + \alpha_i M_t \, \frac{s_{it}}{s_{it} + s_{jt}}.$$

Both channels — direct research and contest winnings — carry the common factor α_i.

Step 2: First-order condition. Since $\partial s_{it} / \partial r_{it} = -\gamma_i$, differentiating with respect to r_{it} (treating s_{jt} as given, as firms move simultaneously) yields

$$\alpha_i \left[1 \;-\; \frac{\gamma_i M_t \, s_{jt}}{(s_{it} + s_{jt})^2} \right] \;=\; 0.$$

The factor $\alpha_i > 0$ cancels, leaving the Tullock condition

$$\frac{\gamma_i M_t \, s_{jt}}{(s_{it} + s_{jt})^2} \;=\; 1. \tag{13}$$

The second-order condition holds because s_{it} is linear in r_{it} and the contest success function is

strictly concave in s_{it} for $s_{jt} > 0$.

Step 3: Equilibrium score ratio. Firm j's analogous first-order condition is

$$\frac{\gamma_j M_t \, s_{it}}{(s_{it} + s_{jt})^2} \;=\; 1.$$

Dividing one by the other gives

$$\frac{s^*_{it}}{s^*_{jt}} \;=\; \frac{\gamma_i}{\gamma_j},$$

independent of the state, the prize, and the horizon. Hence the equilibrium budget share is

$$\frac{B^*_{i,t+1}}{M_t} \;=\; \frac{s^*_{it}}{s^*_{it} + s^*_{jt}} \;=\; \frac{\gamma_i}{\gamma_i + \gamma_j},$$

which matches the budget formula in (5). Substituting $s^*_{it}/s^*_{jt} = \gamma_i/\gamma_j$ back into (13) and solving gives

$$s^*_{it} \;=\; \frac{\gamma_i^2 \gamma_j}{(\gamma_i + \gamma_j)^2} \, M_t,$$

which matches the score formula in (5).

Step 4: Closing the induction. Evaluating the reduced objective at the equilibrium allocation requires expressing r^*_{it} in terms of the state. From $s^*_{it} = \theta Q_{i,t-1} + B_{it} - \gamma_i r^*_{it}$,

$$\gamma_i r^*_{it} \;=\; \theta Q_{i,t-1} + B_{it} - s^*_{it}.$$

Then

$$\begin{aligned} Q^*_{it} + \alpha_i B^*_{i,t+1} &= Q_{i,t-1} + \alpha_i r^*_{it} + \alpha_i \cdot \frac{\gamma_i M_t}{\gamma_i + \gamma_j} \\ &= Q_{i,t-1} + \frac{\alpha_i}{\gamma_i} \left[\theta Q_{i,t-1} + B_{it} - s^*_{it}\right] + \frac{\alpha_i \gamma_i M_t}{\gamma_i + \gamma_j} \\ &= \frac{\gamma_i + \alpha_i \theta}{\gamma_i} \, Q_{i,t-1} + \frac{\alpha_i}{\gamma_i} \, B_{it} - \frac{\alpha_i s^*_{it}}{\gamma_i} + \frac{\alpha_i \gamma_i M_t}{\gamma_i + \gamma_j}. \end{aligned}$$

Using $\gamma_i + \alpha_i \theta = 1$ and substituting $s^*_{it} = \gamma_i^2 \gamma_j M_t/(\gamma_i + \gamma_j)^2$,

$$\begin{aligned} Q^*_{it} + \alpha_i B^*_{i,t+1} &= \frac{Q_{i,t-1} + \alpha_i B_{it}}{\gamma_i} \;-\; \frac{\alpha_i \gamma_i \gamma_j M_t}{(\gamma_i + \gamma_j)^2} \;+\; \frac{\alpha_i \gamma_i M_t}{\gamma_i + \gamma_j} \\ &= \frac{Q_{i,t-1} + \alpha_i B_{it}}{\gamma_i} \;+\; \frac{\alpha_i \gamma_i^2 M_t}{(\gamma_i + \gamma_j)^2}, \end{aligned}$$

where the last equality uses $\gamma_i/(\gamma_i + \gamma_j) - \gamma_i \gamma_j/(\gamma_i + \gamma_j)^2 = \gamma_i^2/(\gamma_i + \gamma_j)^2$. Dividing by γ_i^{T-t-1} and adding D_i^{t+1} recovers (12) at period t:

$$V_i^t(Q_{i,t-1}, B_{it}) \;=\; \frac{Q_{i,t-1} + \alpha_i B_{it}}{\gamma_i^{T-t}} \;+\; D_i^t, \qquad D_i^t \;=\; D_i^{t+1} + \frac{\alpha_i \gamma_i^2 M_t}{\gamma_i^{T-t-1}(\gamma_i + \gamma_j)^2}.$$

This completes the inductive step. Interiority ensures that the candidate score pair is feasible with strictly positive research and hacking for both firms. Uniqueness follows because (13) and its counterpart for firm j admit a unique positive solution (s^*_{it}, s^*_{jt}), and each score uniquely pins down r^*_{it} through the budget identity.

Equilibrium payoff. Evaluating at the initial state $(Q_{i0}, B_{i1}) = (0, B_{i1})$,

$$U_i^* \;=\; V_i^1(0, B_{i1}) \;=\; \frac{\alpha_i B_{i1}}{\gamma_i^{T-1}} + D_i^1.$$

Unwinding the recursion for D_i^t from $D_i^T = 0$ gives

$$D_i^1 \;=\; \sum_{t=1}^{T-1} \frac{\alpha_i \gamma_i^2 M_t}{\gamma_i^{T-t-1}(\gamma_i + \gamma_j)^2} \;=\; \frac{\alpha_i \gamma_i^2}{(\gamma_i + \gamma_j)^2} \sum_{t=1}^{T-1} \frac{M_t}{\gamma_i^{T-1-t}},$$

yielding (6). □

A.2 Proof of Proposition 1

Under the interior equilibrium formulas and equal initial endowments $B_{11} = B_{21} = B_1$, subtracting (6) for firm 1 from the corresponding expression for firm 2 gives

$$U_2^* - U_1^* \;=\; \left(\frac{\alpha_2}{\gamma_2^{T-1}} - \frac{\alpha_1}{\gamma_1^{T-1}}\right) B_1 \;+\; \frac{1}{(\gamma_1 + \gamma_2)^2} \sum_{t=1}^{T-1} \left(\alpha_2 \gamma_2^{t-T+3} - \alpha_1 \gamma_1^{t-T+3}\right) M_t. \tag{14}$$

Therefore the inefficient firm has the higher payoff if and only if

$$\sum_{t=1}^{T-1} \left(\alpha_2 \gamma_2^{t-T+3} - \alpha_1 \gamma_1^{t-T+3}\right) M_t \;>\; (\gamma_1 + \gamma_2)^2 \left(\frac{\alpha_1}{\gamma_1^{T-1}} - \frac{\alpha_2}{\gamma_2^{T-1}}\right) B_1. \tag{15}$$

This is exactly condition (8) in the statement of the proposition, so the result follows immediately. □

A.3 Normalized dynamics under geometric prize growth

Along any interior path, Theorem 1 implies

$$X_{i,t+1} \;=\; \frac{X_{it}}{\gamma_i} \;+\; \frac{\alpha_i \gamma_i^2}{(\gamma_1 + \gamma_2)^2} M_t.$$

If $M_t = M_1 \lambda^{t-1}$ and $x_{it} \equiv X_{it}/M_t$, dividing by $M_{t+1} = \lambda M_t$ yields

$$x_{i,t+1} \;=\; \frac{x_{it}}{\gamma_i \lambda} \;+\; \frac{\alpha_i \gamma_i^2}{(\gamma_1 + \gamma_2)^2 \lambda}.$$

The solution is

$$x_{it} \;=\; \left(\frac{1}{\gamma_i \lambda}\right)^{t-1} x_{i1} \;+\; \frac{\alpha_i \gamma_i^2}{(\gamma_1 + \gamma_2)^2 \lambda} \sum_{k=0}^{t-2} \left(\frac{1}{\gamma_i \lambda}\right)^k.$$

If $\lambda \le 1/\gamma_i$, the coefficient on x_{i1} is weakly increasing in t, and $x_{it} \to \infty$. If $\lambda > 1/\gamma_i$, then $a_i \equiv 1/(\gamma_i \lambda)$ lies in $(0, 1)$, and the recursion has fixed point

$$x_i^\infty \;=\; \frac{\alpha_i \gamma_i^2}{(\gamma_1 + \gamma_2)^2 (\lambda - 1/\gamma_i)}.$$

Moreover,

$$x_{it} - x_i^\infty \;=\; a_i^{t-1}(x_{i1} - x_i^\infty),$$

so x_{it} converges monotonically to x_i^∞. Finally, $\theta X_{it} = \theta x_{it} M_t$ is the truthful component of firm i's leaderboard score, while the interior target score is $s_i^I = \gamma_i^2 \gamma_j M_t/(\gamma_1+\gamma_2)^2$. □

A.4 Proof of Proposition 2

The proposition is proved in three steps. Under constant prizes, the interior target score is fixed while truthful performance keeps accumulating, so the equilibrium must first leave the interior regime. Once truthful scores are high enough, both firms' best responses become full research and an honest tail begins. In that honest tail the efficient firm's true-quality ratio rises monotonically and eventually crosses one, so it overtakes on the truthful leaderboard. From that point onward its true-quality lead grows fast enough to generate finite-time separation.

A.4.1 Exit from the interior regime under constant prizes

Setting $\lambda = 1$ in the normalized recursion above gives $x_{it} = X_{it}/M \to \infty$ for both firms, so X_{it} diverges along any interior path. Under constant prizes, however, the interior target score

$$s_i^I \;=\; \frac{\gamma_i^2 \gamma_j}{(\gamma_1+\gamma_2)^2}\, M$$

is time-invariant. Since interiority requires $\underline{s}_{it} = \theta X_{it} < s_i^I$, there exists a smallest date τ_i such that $\theta X_{i,\tau_i} \geq s_i^I$. Let $\tau_I \equiv \min\{\tau_1, \tau_2\}$. Then for every horizon $T > \tau_I$, the candidate interior score pair from Theorem 1 is infeasible at date τ_I, so the equilibrium path cannot remain interior through every contest date.

It remains to identify which constraint binds. Writing firm i's reduced period payoff in score space,

$$\Phi_i(s_i; s_j) \;=\; C_i \;+\; \frac{\alpha_i}{\gamma_i}\left[-\,s_i \;+\; \gamma_i M \frac{s_i}{s_i+s_j}\right],$$

where C_i is independent of s_i, shows that the objective is strictly concave in s_i. Hence if the unconstrained maximiser lies below the feasible interval, the optimum is the lower endpoint. At date τ_I that is exactly what happens: $s_i^I \leq \underline{s}_{i,\tau_I} = \theta X_{i,\tau_I}$. The binding corner is therefore the lower endpoint, which corresponds to full research and $h^*_{i,\tau_I} = 0$. □

A.4.2 Honest leaderboard tail

Fix a contest date t and let $a_i \equiv \theta X_{it}$ denote firm i's truthful lower-bound leaderboard score. Using the reduced score problem from the proof of Proposition 2,

$$\frac{\partial \Phi_i}{\partial s_i}(a_i; s_j) \;=\; \frac{\alpha_i}{\gamma_i}\left[-1 + \frac{\gamma_i M s_j}{(a_i+s_j)^2}\right].$$

For fixed a_i, the term $s_j/(a_i+s_j)^2$ is maximised at $s_j = a_i$ and its maximum value is $1/(4a_i)$. Therefore if

$$a_i = \theta X_{it} \;\geq\; \frac{\gamma_i M}{4},$$

then $\partial \Phi_i/\partial s_i(a_i; s_j) \leq 0$ for every rival score $s_j \geq 0$. Strict concavity then implies that the unique best response is the lower endpoint, so $h^*_{it} = 0$ regardless of the rival's action.

By the exit argument above, X_{it} grows without bound along any interior path under constant prizes, and once a lower-bound corner binds the firm only shifts further toward research relative to

the infeasible interior target. Hence there exists a finite date $\tau_H \geq \tau_I$ such that

$$\theta X_{1t} \geq \frac{\gamma_1 M}{4} \qquad \text{and} \qquad \theta X_{2t} \geq \frac{\gamma_2 M}{4} \qquad \text{for all } t \geq \tau_H.$$

From that date onward both firms choose $h_{it}^* = 0$, so the tail is honest. □

A.4.3 Eventual truthful overtaking in the honest tail

Fix the honest-tail date τ_H defined in the main text. For every $t \geq \tau_H$, hacking is zero, so

$$s_{it}^* = \theta Q_{it} \qquad \text{and} \qquad B_{i,t+1}^* = M \frac{Q_{it}}{Q_{1t} + Q_{2t}}.$$

Hence true quality evolves according to

$$Q_{i,t+1} \;=\; Q_{it} + \alpha_i M \frac{Q_{it}}{Q_{1t} + Q_{2t}}, \qquad t \geq \tau_H.$$

Define the true-quality ratio $R_t \equiv Q_{1t}/Q_{2t}$. Then

$$R_{t+1} \;=\; R_t \frac{1 + \alpha_1 M/(Q_{1t} + Q_{2t})}{1 + \alpha_2 M/(Q_{1t} + Q_{2t})} \;>\; R_t,$$

because $\alpha_1 > \alpha_2$. So R_t is strictly increasing throughout the honest tail.

It remains to show that this increase is large enough to imply eventual overtaking. Let $S_t \equiv Q_{1t} + Q_{2t}$. Since $B_{1,t+1}^* + B_{2,t+1}^* = M$ and $\alpha_1 > \alpha_2$,

$$S_{t+1} \;=\; S_t + \alpha_1 B_{1,t+1}^* + \alpha_2 B_{2,t+1}^* \;\leq\; S_t + \alpha_1 M.$$

Therefore S_t grows at most linearly. Also,

$$\log R_{t+1} - \log R_t \;=\; \log\!\left(1 + \frac{(\alpha_1 - \alpha_2)M}{S_t + \alpha_2 M}\right).$$

Using $\log(1+u) \geq u/(1+u)$ for $u \geq 0$ gives

$$\log R_{t+1} - \log R_t \;\geq\; \frac{(\alpha_1 - \alpha_2)M}{S_t + \alpha_1 M}.$$

Since the denominator on the right-hand side is at most linear in t, the series of lower bounds diverges. Hence $\log R_t \to \infty$, so $R_t \to \infty$. There is therefore a finite date $\tau_O \geq \tau_H$ such that

$$Q_{1t} \geq Q_{2t} \qquad \text{for all } t \geq \tau_O.$$

Because the leaderboard is truthful on the honest tail, this means firm 1 weakly overtakes firm 2 on the leaderboard from date τ_O onward. □

A.4.4 Finite-time leaderboard separation

Fix the truthful-overtake date τ_O established above. Since the leaderboard is truthful from τ_H onward and $Q_{1t} \geq Q_{2t}$ for all $t \geq \tau_O$, it follows that for every $t \geq \tau_O + 1$,

$$B_{1t}^* \geq \frac{M}{2}, \qquad B_{2t}^* \leq \frac{M}{2}.$$

Define the truthful quality gap $D_t \equiv Q_{1t} - Q_{2t}$. Along the honest tail,

$$D_t \;=\; D_{t-1} + \alpha_1 B^*_{1t} - \alpha_2 B^*_{2t} \;\geq\; D_{t-1} + \frac{M}{2}(\alpha_1 - \alpha_2), \qquad t \geq \tau_O + 1.$$

Since $\alpha_1 > \alpha_2$, the gap grows by at least $M(\alpha_1 - \alpha_2)/2 > 0$ every period. Iterating,

$$D_t \;\geq\; D_{\tau_O} + (t - \tau_O)\frac{M}{2}(\alpha_1 - \alpha_2), \qquad t \geq \tau_O,$$

so $D_t \to \infty$ linearly.

Now compare firm 1's truthful lower-bound score to firm 2's maximal feasible score:

$$\begin{aligned}
\underline{s}_{1t} - \overline{s}_{2t} &= \big(\theta Q_{1,t-1} + \theta\alpha_1 B^*_{1t}\big) - \big(\theta Q_{2,t-1} + B^*_{2t}\big) \\
&= \theta D_{t-1} + \theta\alpha_1 B^*_{1t} - B^*_{2t} \\
&\geq \theta D_{t-1} + \theta\alpha_1 \frac{M}{2} - \frac{M}{2} \\
&= \theta D_{t-1} - \frac{\gamma_1 M}{2}.
\end{aligned}$$

Because $D_{t-1} \to \infty$, there exists a finite date $\tau_S \geq \tau_H$ such that

$$\theta D_{t-1} > \frac{\gamma_1 M}{2} \qquad \text{for all } t \geq \tau_S.$$

For every such t,

$$\underline{s}_{1t} - \overline{s}_{2t} > 0,$$

that is, $\underline{s}_{1t} > \overline{s}_{2t}$. Therefore from date τ_S onward, even full hacking by firm 2 cannot overturn firm 1's truthful leaderboard lead. □

A.4.5 Bound on the time to leaderboard separation

The separation argument just established shows that, for every $t \geq \tau_O$,

$$D_t \;\geq\; D_{\tau_O} + (t - \tau_O)\frac{M}{2}(\alpha_1 - \alpha_2),$$

and that separation occurs whenever $D_{t-1} > \gamma_1 M/(2\theta)$. If n_S is the smallest integer satisfying

$$D_{\tau_O} + n_S \frac{M}{2}(\alpha_1 - \alpha_2) > \frac{\gamma_1 M}{2\theta},$$

then the threshold is crossed by contest date $t = \tau_O + 1 + n_S$. Therefore $\tau_S \leq \tau_O + 1 + n_S$. □

A.5 Proof of Proposition 3

Define $x_{it} \equiv X_{it}/M_t$. By the normalized recursion above, the lower bound on λ implies $1/(\gamma_i\lambda) \in (0, 1)$ for both firms, and

$$x_{it} - x_i^\infty \;=\; \left(\frac{1}{\gamma_i\lambda}\right)^{t-1}(x_{i1} - x_i^\infty), \qquad x_i^\infty \;=\; \frac{\alpha_i\gamma_i^2}{(\gamma_1 + \gamma_2)^2(\lambda - 1/\gamma_i)}.$$

So x_{it} moves monotonically from x_{i1} to x_i^∞.

Along any interior path, lower feasibility at contest date t is

$$\theta x_{it} \;<\; \frac{\gamma_i^2\gamma_j}{(\gamma_1+\gamma_2)^2}. \tag{LF}$$

The first contest satisfies (LF) by assumption. Since x_{it} stays between x_{i1} and x_i^∞, it is enough that (LF) hold at the limit. The limit inequality

$$\theta x_i^\infty \;<\; \frac{\gamma_i^2\gamma_j}{(\gamma_1+\gamma_2)^2}$$

is equivalent to

$$\lambda \;>\; \frac{1}{\gamma_i}+\frac{\theta\alpha_i}{\gamma_j}.$$

For $i=1$ this is the stated lower bound. It is tighter than the corresponding condition for $i=2$ because

$$\frac{1}{\gamma_i}+\frac{\theta\alpha_i}{\gamma_j} \;=\; \frac{1}{\gamma_1}+\frac{1}{\gamma_2}-\frac{\gamma_i}{\gamma_j},$$

and $\gamma_1/\gamma_2 < \gamma_2/\gamma_1$.

For upper feasibility, first-contest interiority handles $t=1$. Fix $t\geq 2$. If contest $t-1$ is interior, then

$$B_{it} \;=\; M_{t-1}\frac{\gamma_i}{\gamma_1+\gamma_2} \;=\; M_t\frac{\gamma_i}{(\gamma_1+\gamma_2)\lambda},$$

so

$$\frac{\overline{s}_{it}}{M_t} \;=\; \theta x_{it}+\frac{\gamma_i^2}{(\gamma_1+\gamma_2)\lambda}. \tag{UF}$$

By first-contest interiority,

$$B_1 \;>\; \frac{\gamma_i^2\gamma_j}{(\gamma_1+\gamma_2)^2}M_1,$$

so $x_{i1}=\alpha_iB_1/M_1 > \alpha_i\gamma_i^2\gamma_j/(\gamma_1+\gamma_2)^2$. Using the recursion once,

$$\begin{aligned}\theta x_{i2}+\frac{\gamma_i^2}{(\gamma_1+\gamma_2)\lambda} &> \frac{\theta\alpha_i\gamma_i\gamma_j+\theta\alpha_i\gamma_i^2+\gamma_i^2(\gamma_1+\gamma_2)}{(\gamma_1+\gamma_2)^2\lambda}\\ &= \frac{\gamma_i}{(\gamma_1+\gamma_2)\lambda},\end{aligned}$$

where the last equality uses $\theta\alpha_i+\gamma_i=1$. Therefore the upper-feasibility inequality at $t=2$,

$$\frac{\gamma_i^2\gamma_j}{(\gamma_1+\gamma_2)^2} \;<\; \theta x_{i2}+\frac{\gamma_i^2}{(\gamma_1+\gamma_2)\lambda},$$

holds whenever

$$\lambda \;<\; \frac{\gamma_1+\gamma_2}{\gamma_i\gamma_j} \;=\; \frac{1}{\gamma_i}+\frac{1}{\gamma_j}.$$

The expression on the right-hand side of (UF) is monotone in t because x_{it} is monotone. Its limit is

$$\theta x_i^\infty+\frac{\gamma_i^2}{(\gamma_1+\gamma_2)\lambda},$$

and the inequality

$$\frac{\gamma_i^2\gamma_j}{(\gamma_1+\gamma_2)^2} \;<\; \theta x_i^{\infty} + \frac{\gamma_i^2}{(\gamma_1+\gamma_2)\lambda}$$

is again equivalent to $\lambda < 1/\gamma_i + 1/\gamma_j$. So upper feasibility holds for every $t \geq 2$.

Since contest 1 is interior by assumption and the same feasibility inequalities apply at any later date once the previous contest is interior, induction implies that the path is interior at every contest date. Theorem 1 then gives

$$\frac{s_{2t}^*}{s_{1t}^*} \;=\; \frac{\gamma_2}{\gamma_1} \;>\; 1, \qquad \frac{B_{2,t+1}^*}{B_{1,t+1}^*} \;=\; \frac{\gamma_2}{\gamma_1} \;>\; 1$$

for all t. □

A.6 Proof of Proposition 5

Fix $\bar{M} = \sum_{t=1}^{T-1} M_t$ and consider only prize sequences whose equilibrium path is interior at every contest date. The welfare identity (9) gives

$$W^{\mathrm{LB}} \;=\; (\alpha_1+\alpha_2)B_1 \;+\; \frac{\alpha_1\gamma_1+\alpha_2\gamma_2}{\gamma_1+\gamma_2}\,\bar{M} \;-\; \sum_{\tau=1}^{T-1}\Delta_\tau,$$

so under fixed $\bar{M}$, welfare differences across interior prize sequences are determined entirely by $-\sum_\tau \Delta_\tau$. We show that the marginal effect of prize mass on total hacking is strictly larger at later dates; the transfer claim in the proposition then follows.

Step 1: Closed form for the equilibrium state. Define $X_{it} \equiv Q_{i,t-1} + \alpha_i B_{it}$, the scalar state entering period t. From the algebra in the proof of Theorem 1 (see the transition at (12)), the equilibrium allocation satisfies the linear recursion

$$X_{i,t+1} \;=\; \frac{X_{it}}{\gamma_i} \;+\; \frac{\alpha_i\gamma_i^2\,M_t}{(\gamma_1+\gamma_2)^2}, \qquad X_{i1} \;=\; \alpha_i B_1.$$

Iterating,

$$X_{it} \;=\; \frac{\alpha_i B_1}{\gamma_i^{t-1}} \;+\; \frac{\alpha_i\gamma_i^2}{(\gamma_1+\gamma_2)^2}\sum_{\sigma=1}^{t-1}\frac{M_\sigma}{\gamma_i^{t-1-\sigma}}.$$

Hence $\partial X_{it}/\partial M_\sigma = \alpha_i\gamma_i^{\sigma-t+3}/(\gamma_1+\gamma_2)^2$ for $\sigma < t$ and zero otherwise.

Step 2: Closed form for equilibrium hacking. Combining the budget identity $h_{it}^* = B_{it} - r_{it}^*$ with the first-order condition $\gamma_i r_{it}^* = \theta Q_{i,t-1} + B_{it} - s_{it}^*$ and using $\gamma_i + \theta\alpha_i = 1$,

$$h_{it}^* \;=\; \frac{s_{it}^* - \theta X_{it}}{\gamma_i}, \qquad s_{it}^* \;=\; \frac{\gamma_i^2\gamma_j\,M_t}{(\gamma_1+\gamma_2)^2}.$$

Crucially, s_{it}^* depends only on M_t (not on any earlier prize) and X_{it} depends only on $\{M_\sigma\}_{\sigma<t}$.

Step 3: Cross-period comparative statics. Differentiating $\alpha_i h_{it}^*$,

$$\alpha_i\,\frac{\partial h_{it}^*}{\partial M_\tau} \;=\; \begin{cases} \dfrac{\alpha_i\gamma_i\gamma_j}{(\gamma_1+\gamma_2)^2} & \text{if } \tau = t, \\[2ex] -\dfrac{\theta\alpha_i^2\,\gamma_i^{\tau-t+2}}{(\gamma_1+\gamma_2)^2} & \text{if } \tau < t, \\[2ex] 0 & \text{if } \tau > t. \end{cases}$$

Summing across firms, $\partial\Delta_t/\partial M_t = \gamma_1\gamma_2(\alpha_1+\alpha_2)/(\gamma_1+\gamma_2)^2 > 0$ and $\partial\Delta_\tau/\partial M_t = -\theta(\alpha_1^2\gamma_1^{t-\tau+2} + \alpha_2^2\gamma_2^{t-\tau+2})/(\gamma_1+\gamma_2)^2 < 0$ for $\tau > t$, with zero effect for $\tau < t$.

Step 4: Ranking the prize periods. Let

$$G_t \;\equiv\; \frac{\partial\sum_\tau \Delta_\tau}{\partial M_t}.$$

By Step 3,

$$G_t \;=\; \frac{\gamma_1\gamma_2(\alpha_1+\alpha_2)}{(\gamma_1+\gamma_2)^2} \;-\; \frac{\theta}{(\gamma_1+\gamma_2)^2}\sum_{i=1,2}\alpha_i^2\sum_{k=1}^{T-1-t}\gamma_i^{2-k},$$

where the substitution $k = \tau - t$ re-indexes the future-period terms. The "self-hacking" term on the left is the same for every t; only the future-dampening sum depends on t, and it is strictly decreasing in t because the upper limit $T-1-t$ falls as t rises and every term γ_i^{2-k} is strictly positive. Therefore for $t < t'$,

$$G_t - G_{t'} \;=\; -\frac{\theta}{(\gamma_1+\gamma_2)^2}\sum_{i=1,2}\alpha_i^2\sum_{k=T-t'}^{T-1-t}\gamma_i^{2-k} \;<\; 0.$$

Since the derivatives in Steps 1–3 depend only on primitives and the horizon, G_t is constant on the interior region. Hence, if $\tilde{M}$ is obtained from M by moving ε units of prize mass from t' to t, then

$$\sum_\tau \Delta_\tau(\tilde{M}) - \sum_\tau \Delta_\tau(M) \;=\; \varepsilon\,(G_t - G_{t'}) \;<\; 0.$$

Therefore $W^{\mathrm{LB}}(\tilde{M}) > W^{\mathrm{LB}}(M)$. The monotonicity $M_1^* \geq \cdots \geq M_{T-1}^*$ of any welfare-maximising interior prize sequence follows because any inversion $M_t^* < M_{t+1}^*$ would be improved by transferring a small amount of prize mass from $t+1$ to t while staying in the interior region. □

A.7 Proof of Proposition 4

Define

$$\delta_t \;\equiv\; M_t(\alpha_1-\alpha_2)\left(\sigma_{1t}^{\mathrm{LB}} - \frac{1}{2}\right) - \Delta_t^{\mathrm{LB}}, \qquad W^{\mathrm{LB}}(T) - W^{\mathrm{RND}}(T) \;=\; \sum_{t=1}^{T-1}\delta_t.$$

Part (i): flat prizes. Fix $M_t \equiv M$. Because the first contest date is interior, Theorem 1 gives $\sigma_{11}^{\mathrm{LB}} = \gamma_1/(\gamma_1+\gamma_2) < 1/2$ and $\Delta_1^{\mathrm{LB}} > 0$, so $\delta_1 < 0$. Hence $T^- = 2$ works.

By Proposition 2, there exists a finite truthful-overtake date τ_O such that the leaderboard is honest from some earlier finite date onward and firm 1 weakly leads on that truthful leaderboard for every contest $t \geq \tau_O$. On the honest tail, $\Delta_t^{\mathrm{LB}} = 0$ and

$$\sigma_{1t}^{\mathrm{LB}} \;=\; \frac{Q_{1t}}{Q_{1t}+Q_{2t}}.$$

The proof of Proposition 2 shows that $Q_{1t}/Q_{2t} \to \infty$ on that tail, so $\sigma_{1t}^{\mathrm{LB}} \to 1$. Therefore there exists a finite date $\tau_S \geq \tau_O$ such that $\sigma_{1t}^{\mathrm{LB}} \geq 3/4$ for all $t \geq \tau_S$. For those dates,

$$\delta_t \;=\; M(\alpha_1-\alpha_2)\left(\sigma_{1t}^{\mathrm{LB}} - \frac{1}{2}\right) \;\geq\; \frac{M(\alpha_1-\alpha_2)}{4} \;>\; 0.$$

Let $A \equiv \sum_{t=1}^{\tau_S - 1} \delta_t$, which is finite. Then for every $T \geq \tau_S + 1$,

$$W^{\mathrm{LB}}(T) - W^{\mathrm{RND}}(T) \;\geq\; A + (T - \tau_S)\frac{M(\alpha_1 - \alpha_2)}{4}.$$

The right-hand side diverges to $+\infty$, so there exists T^+ such that $W^{\mathrm{LB}}(T) > W^{\mathrm{RND}}(T)$ for all $T \geq T^+$.

Part (ii): fast-growing prizes. Now let $M_t = M_1 \lambda^{t-1}$ with λ in the perpetual-deception region. Proposition 3 implies that the equilibrium path is interior at every contest date. Hence $\sigma_{1t}^{\mathrm{LB}} = \gamma_1/(\gamma_1 + \gamma_2) < 1/2$ and $\Delta_t^{\mathrm{LB}} > 0$ for every t. Therefore

$$\delta_t \;=\; M_t(\alpha_1 - \alpha_2)\frac{\gamma_1 - \gamma_2}{2(\gamma_1 + \gamma_2)} - \Delta_t^{\mathrm{LB}} \;<\; 0 \qquad \text{for every } t.$$

So $W^{\mathrm{LB}}(T) < W^{\mathrm{RND}}(T)$ for every $T \geq 2$.

Moreover, since $\lambda > 1/\gamma_1 + \theta\alpha_1/\gamma_2 > 1$, the geometric prize sequence satisfies $\sum_{t=1}^{T-1} M_t \to \infty$. Let

$$c \;\equiv\; \frac{(\alpha_1 - \alpha_2)(\gamma_2 - \gamma_1)}{2(\gamma_1 + \gamma_2)} \;>\; 0.$$

Then $\delta_t \leq -cM_t$ for every t, so

$$W^{\mathrm{RND}}(T) - W^{\mathrm{LB}}(T) \;\geq\; c\sum_{t=1}^{T-1} M_t \;\to\; \infty.$$

This proves the claim. □

A.8 Sequential leaderboard characterization

Fix a mover order with leader L and follower F in every contest period. The proof again proceeds by backward induction. We show that in any continuation subgame the value function remains affine in the composite state $Q + \alpha_i B$:

$$V_i^t(Q, B) \;=\; \frac{Q + \alpha_i B}{\gamma_i^{T-t}} + D_{i,\mathrm{seq}}^t,$$

where $D_{i,\mathrm{seq}}^t$ depends on the remaining prizes and the mover order, but not on the state.

Base case ($t = T$). The terminal period is unchanged: firms invest their entire budgets in research, so

$$V_i^T(Q, B) \;=\; Q + \alpha_i B.$$

Inductive step. Suppose the affine form holds at $t + 1$. In contest period $t < T$, write choices in terms of posted score rather than research. Using $s_{it} = \theta Q_{i,t-1} + \theta\alpha_i r_{it} + h_{it}$ with $r_{it} + h_{it} = B_{it}$ gives

$$s_{it} \;=\; \theta Q_{i,t-1} + \theta\alpha_i B_{it} + \gamma_i h_{it}, \qquad h_{it} \;=\; \frac{s_{it} - \theta Q_{i,t-1} - \theta\alpha_i B_{it}}{\gamma_i}.$$

Hence

$$Q_{it} \;=\; Q_{i,t-1} + \alpha_i(B_{it} - h_{it}) \;=\; \frac{Q_{i,t-1} + \alpha_i B_{it} - \alpha_i s_{it}}{\gamma_i}.$$

Substituting into the continuation value at $t+1$,

$$V_i^{t+1}(Q_{it}, B_{i,t+1}) \;=\; \frac{Q_{i,t-1} + \alpha_i B_{it}}{\gamma_i^{T-t}} + \frac{\alpha_i}{\gamma_i^{T-t-1}} \left[M_t \frac{s_{it}}{s_{it} + s_{jt}} - \frac{s_{it}}{\gamma_i} \right] + D_{i,\text{seq}}^{t+1}.$$

So, up to a positive multiplicative constant and a state-dependent constant, every contest subgame reduces to the static score-choice problem on the feasible score interval,

$$\max_{s_i \geq 0} \left[M_t \frac{s_i}{s_i + s_j} - \frac{s_i}{\gamma_i} \right].$$

Follower. Given a leader score $\hat{s}_t > 0$, the follower solves

$$\max_{s_F \geq 0} \left[M_t \frac{s_F}{s_F + \hat{s}_t} - \frac{s_F}{\gamma_F} \right].$$

The first-order condition is

$$M_t \frac{\hat{s}_t}{(s_F + \hat{s}_t)^2} \;=\; \frac{1}{\gamma_F},$$

so the follower best response is

$$s_F(\hat{s}_t) \;=\; \sqrt{M_t \gamma_F \hat{s}_t} - \hat{s}_t.$$

It is strictly positive if and only if $\hat{s}_t < M_t \gamma_F$.

Leader. Anticipating the follower best response, the leader solves

$$\max_{\hat{s}_t \geq 0} \left[M_t \sqrt{\frac{\hat{s}_t}{M_t \gamma_F}} - \frac{\hat{s}_t}{\gamma_L} \right].$$

The first-order condition yields

$$\hat{s}_t^* \;=\; \frac{\gamma_L^2}{4\gamma_F} M_t.$$

Substituting back into the follower best response gives

$$s_{Ft}^* \;=\; \frac{\gamma_L(2\gamma_F - \gamma_L)}{4\gamma_F} M_t.$$

The follower score is strictly positive if and only if $\gamma_L < 2\gamma_F$, which is exactly the follower-participation condition. The corresponding budget shares are

$$B_{L,t+1}^* \;=\; M_t \frac{s_{Lt}^*}{s_{Lt}^* + s_{Ft}^*} \;=\; \frac{\gamma_L}{2\gamma_F} M_t, \qquad B_{F,t+1}^* \;=\; \frac{2\gamma_F - \gamma_L}{2\gamma_F} M_t.$$

This gives the Stackelberg score and budget formulas used in Section 6.

Closing the induction. Evaluating $Q_{it} + \alpha_i B_{i,t+1}$ at the equilibrium actions yields, for the leader,

$$Q_{Lt}^* + \alpha_L B_{L,t+1}^* \;=\; \frac{Q_{L,t-1} + \alpha_L B_{Lt}}{\gamma_L} + \alpha_L \frac{\gamma_L}{4\gamma_F} M_t,$$

and for the follower,

$$Q^*_{Ft} + \alpha_F B^*_{F,t+1} \;=\; \frac{Q_{F,t-1} + \alpha_F B_{Ft}}{\gamma_F} + \alpha_F \frac{(2\gamma_F - \gamma_L)^2}{4\gamma_F^2}\, M_t.$$

Dividing by γ_i^{T-t-1} and adding $D^{t+1}_{i,\text{seq}}$ restores the affine form at period t, with new constants

$$D^t_{L,\text{seq}} \;=\; D^{t+1}_{L,\text{seq}} + \frac{\alpha_L \gamma_L}{4\gamma_F}\, \frac{M_t}{\gamma_L^{T-t-1}}, \qquad D^t_{F,\text{seq}} \;=\; D^{t+1}_{F,\text{seq}} + \frac{\alpha_F (2\gamma_F - \gamma_L)^2}{4\gamma_F^2}\, \frac{M_t}{\gamma_F^{T-t-1}}.$$

Finally, $s^*_{it} = \theta Q_{i,t-1} + B_{it} - \gamma_i r^*_{it}$ implies

$$r^*_{it} \;=\; \frac{\theta Q_{i,t-1} + B_{it} - s^*_{it}}{\gamma_i}, \qquad h^*_{it} \;=\; B_{it} - r^*_{it}.$$

Interiority guarantees these actions are strictly feasible. Uniqueness follows because the follower best response is unique and the leader's objective is strictly concave in its posted score on the interior. □

Comparisons used in the text. In the symmetric case $\gamma_1 = \gamma_2 \equiv \gamma$, the sequential characterization gives $s_L = s_F = \gamma M_t/4$ and $B_{L,t+1} = B_{F,t+1} = M_t/2$. With equal initial conditions, symmetry is therefore preserved and the sequential leaderboard coincides exactly with the symmetric simultaneous leaderboard.

Now suppose $\alpha_1 > \alpha_2$, so $\gamma_1 < \gamma_2$. If firm 1 moves first, then

$$s^{\text{eff first}}_{2t} - s^{\text{eff first}}_{1t} \;=\; \frac{\gamma_1(\gamma_2 - \gamma_1)}{2\gamma_2} M_t, \qquad B^{\text{eff first}}_{2,t+1} - B^{\text{eff first}}_{1,t+1} \;=\; \frac{\gamma_2 - \gamma_1}{\gamma_2} M_t.$$

Relative to the simultaneous leaderboard, firm 2's budget share rises because

$$\frac{2\gamma_2 - \gamma_1}{2\gamma_2} - \frac{\gamma_2}{\gamma_1 + \gamma_2} \;=\; \frac{\gamma_1(\gamma_2 - \gamma_1)}{2\gamma_2(\gamma_1 + \gamma_2)} \;>\; 0.$$

If instead firm 2 moves first, and $\gamma_2 < 2\gamma_1$, then

$$s^{\text{ineff first}}_{2t} - s^{\text{ineff first}}_{1t} \;=\; \frac{\gamma_2(\gamma_2 - \gamma_1)}{2\gamma_1} M_t, \qquad B^{\text{ineff first}}_{2,t+1} - B^{\text{ineff first}}_{1,t+1} \;=\; \frac{\gamma_2 - \gamma_1}{\gamma_1} M_t.$$

Moreover,

$$\frac{\gamma_2}{2\gamma_1} - \frac{2\gamma_2 - \gamma_1}{2\gamma_2} \;=\; \frac{(\gamma_2 - \gamma_1)^2}{2\gamma_1\gamma_2} \;>\; 0,$$

so the inefficient firm's funding share is largest when it moves first.

For the individual hacking comparisons, fix a continuation state and let $h_i^{\text{sym}(i)}$ denote firm i's hacking in the symmetric contest against an identical rival with the same prize and state, so the target score is $\gamma_i M_t/4$. Since $h_i = (s_i - \theta Q_{i,t-1} - \theta\alpha_i B_{it})/\gamma_i$, differences in hacking are differences in target scores scaled by $1/\gamma_i$. Thus, if the efficient firm leads,

$$h^{\text{eff first}}_{1t} - h^{\text{sym}(1)}_{1t} \;=\; -\frac{\gamma_2 - \gamma_1}{4\gamma_2} M_t, \qquad h^{\text{eff first}}_{2t} - h^{\text{sym}(2)}_{2t} \;=\; -\frac{(\gamma_2 - \gamma_1)^2}{4\gamma_2^2} M_t,$$

while if the inefficient firm leads,

$$h^{\text{ineff first}}_{2t} - h^{\text{sym}(2)}_{2t} \;=\; \frac{\gamma_2 - \gamma_1}{4\gamma_1} M_t, \qquad h^{\text{ineff first}}_{1t} - h^{\text{sym}(1)}_{1t} \;=\; -\frac{(\gamma_2 - \gamma_1)^2}{4\gamma_1^2} M_t.$$

For total hacking, fix a continuation state. Since

$$h_{1t} + h_{2t} \;=\; \left(\frac{s_{1t}}{\gamma_1} + \frac{s_{2t}}{\gamma_2}\right) - \theta\left(\frac{Q_{1,t-1} + \alpha_1 B_{1t}}{\gamma_1} + \frac{Q_{2,t-1} + \alpha_2 B_{2t}}{\gamma_2}\right),$$

the ranking of hacking is the ranking of the weighted score target. Under efficient-first sequential play,

$$\left(\frac{s_{1t}}{\gamma_1} + \frac{s_{2t}}{\gamma_2}\right)^{\text{eff first}} = \frac{\gamma_1(3\gamma_2 - \gamma_1)}{4\gamma_2^2} M_t,$$

while under simultaneous play it is $2\gamma_1\gamma_2 M_t/(\gamma_1+\gamma_2)^2$, and the difference is

$$\frac{2\gamma_1\gamma_2}{(\gamma_1+\gamma_2)^2} M_t - \frac{\gamma_1(3\gamma_2-\gamma_1)}{4\gamma_2^2} M_t \;=\; \frac{\gamma_1(\gamma_2-\gamma_1)(5\gamma_2^2-\gamma_1^2)}{4\gamma_2^2(\gamma_1+\gamma_2)^2} M_t \;>\; 0.$$

Under inefficient-first sequential play,

$$\left(\frac{s_{1t}}{\gamma_1} + \frac{s_{2t}}{\gamma_2}\right)^{\text{ineff first}} = \frac{\gamma_2(3\gamma_1 - \gamma_2)}{4\gamma_1^2} M_t,$$

and, on the interior region $\gamma_2 < 2\gamma_1$,

$$\frac{\gamma_2(3\gamma_1-\gamma_2)}{4\gamma_1^2} M_t - \frac{2\gamma_1\gamma_2}{(\gamma_1+\gamma_2)^2} M_t \;=\; \frac{\gamma_2(\gamma_2-\gamma_1)(5\gamma_1^2-\gamma_2^2)}{4\gamma_1^2(\gamma_1+\gamma_2)^2} M_t \;>\; 0.$$

So efficient-first sequentiality lowers contemporaneous hacking relative to simultaneous play, while inefficient-first sequentiality raises it.

B Additional Empirical Details

B.1 Data sources and sample construction

Leaderboard rankings. The v1 and v2 rankings are reconstructed directly from the Hugging Face datasets that back the public leaderboards: `open-llm-leaderboard-old/contents` for v1 and `open-llm-leaderboard/contents` for v2. For v1 I retain rows flagged `Maintainers Choice == True` (the curated subset shown on the live v1 leaderboard) and deduplicate by model `fullname`, keeping the highest-scoring precision variant (313 models). For v2 I retain rows flagged `Official Providers == True` with `Submission Date` on or before 2024-06-30 (the launch-week cohort), then deduplicate analogously (218 models). The top of the reconstructed v2 leaderboard matches the official launch announcement exactly (Qwen2-72B-Instruct #1, Meta-Llama-3-70B-Instruct #2, etc.). The intersection of the two filtered sets — models appearing on both leaderboards, so that ΔRank_i is defined — is 145 models; the main event-study sample is 144 of these with at least one observed download snapshot.

Pre-shock downloads. Monthly downloads are measured at three pre-shock dates drawn from independent sources:

- **October 2023:** a local SQLite snapshot from the PeaTMOSS research dataset, representing Hugging Face model download counts collected in late October 2023 (111 matches in the curated leaderboard intersection).
- **March 5, 2024:** the Hugging Face dataset `SKSowe/Models_Downloads_and_Likes_Metrics` (101 matches in the curated leaderboard intersection).

- **April 17, 2024:** the Kaggle dataset `bwandowando/huggingface-co-model-catalogue`, a broad catalogue pull of Hugging Face models (116 matches in the curated leaderboard intersection).

All three are drawn from Hugging Face's standard 30-day rolling monthly download metric, so cross-source comparisons measure the same underlying quantity.

Post-shock downloads. In the main panel, post-shock downloads are measured at seven pinned snapshots of `cfahlgren1/hub-stats`, a commit-indexed archive of Hugging Face's public hub statistics, taken on August 8, August 19, August 27, September 27, and December 27 of 2024, and March 27 and June 27 of 2025. The July 24, 2024 snapshot, whose 30-day rolling window still overlaps the pre-shock period, is excluded from the event-study panel. A secondary Wayback Machine scrape of archived Hugging Face model pages dated May–June 2024 provides an independent pre-period measurement used for robustness (220 captures, 209 clean matches).

B.2 Variable definitions

For each model i in the intersection sample I compute

$$\Delta\text{Rank}_i \;=\; \text{rank}_i^{v1} - \text{rank}_i^{v2}, \qquad \ln(1+\text{dl}_{it}),$$

where ΔRank_i is positive when the model improved in rank between v1 and v2, and dl_{it} is the raw 30-day download count at date t. For the two-period first-difference exercise reported in the main text, the post-shock outcome $\overline{\text{dl}}_i^{\text{post}}$ averages the December 2024, March 2025, and June 2025 snapshots; this averaging smooths the month-to-month noise inherent in a rolling 30-day count and gives a compact measure of persistent demand rather than a one-month fluctuation. The first-difference outcome is then

$$\Delta \ln(\text{dl})_i \;=\; \ln\!\left(1+\overline{\text{dl}}_i^{\text{post}}\right) \;-\; \ln(1+\text{dl}_{i,\,\text{Apr 17}})\,.$$

For the ANCOVA robustness, $\ln(\text{age}_i)$ is the log number of days between the model's `created_at` date and June 27, 2024, and model-type fixed effects group models into four coarse categories: chat, pretrained, fine-tuned, and merge. The ANCOVA sample falls from 116 to 98 models because eighteen observations in the long-run first-difference sample lack a usable model-type classification.

B.3 Robustness: alternative specifications

Table 1 collects the main estimate under several alternative estimators and sample-sensitivity tests, all on the long-run averaged post-shock outcome.

The point estimate survives replacing the first-difference with a lagged-dependent-variable ANCOVA, winsorising either the outcome or the treatment, removing the five most influential observations, bootstrap inference, clustering standard errors at the coarse model-type level, replacing the April pre-period with a Wayback-based measurement, and re-estimating the same relationship in the panel event-study.

B.4 Robustness: alternative functional forms of ΔRank

Table 2 reports the first-difference estimate at the June 2025 horizon under alternative transformations of the rank-change treatment.

Specification	$\hat{\beta}$	**SE**	p
Baseline FD (main text)	+0.0101	0.0033	0.002
ANCOVA with $\ln(\text{dl}_{\text{pre}})$, ln(age), model-type FE	+0.0126	0.0044	0.004
FD, outcome winsorised at 5/95	+0.0081	0.0036	0.024
FD, treatment ΔRank winsorised at 5/95	+0.0113	–	0.003
ANCOVA, drop 5 highest-Cook's-distance observations	+0.0100	–	< 0.001
Bootstrap (1,000 resamples)	+0.0100	95% CI [+0.003, +0.016]	
Cluster SE by model-type (5 clusters)	+0.0107	–	0.06
Wayback pre-period (March 2025 horizon)	+0.0121	–	0.020
TWFE event study, pooled post (equation (1))	+0.0082	0.0031	0.008
Unfiltered v1∩v2 intersection, ANCOVA	+0.00022	–	0.014

Table 1: Alternative specifications for the relation between ΔRank_i and log downloads. Except where noted, the sample is the launch-cohort Official-Providers intersection, with estimating samples ranging from $N = 98$ to $N = 144$ depending on outcome availability and ancillary controls; the baseline long-run outcome averages the December 2024, March 2025, and June 2025 snapshots. Standard errors are HC3 robust in the FD and ANCOVA rows and clustered at the model level in the TWFE row. The unfiltered row replaces the curated filters with a deduplication-only filter; see Appendix B.5.

Treatment transform	$\hat{\beta}$	**SE**	p
Linear ΔRank	+0.0132	0.0040	0.001
Percentile $\Delta\text{Rank}/N$	+3.99	1.43	0.005
$\Delta \ln(\text{rank})$	+0.698	0.41	0.086
$\Delta(1/\text{rank})$	+11.4	12.4	0.36
Quadratic $\Delta\text{Rank}^2 \cdot \text{sign}$	+0.00004	0.0001	0.71
Entered v2 top-20 indicator	−0.21	1.27	0.87
Entered v2 top-50 indicator	−0.28	0.55	0.61
PPML on levels (offset = ln pre)	+0.0063	0.0012	$< 10^{-7}$

Table 2: Alternative functional forms of the rank-change treatment, evaluated at the June 2025 horizon. Linear, percentile-rank, and PPML are all positive and significant at conventional levels. Threshold-crossing indicators lack power in this sample (≤ 3 intersection models cross the v2 top-20 threshold).

B.5 Unfiltered intersection

Replacing the Maintainers-Choice and Official-Providers filters with a deduplication-only filter enlarges the intersection from 145 to 397 models, but changes the meaning of ΔRank_i. The unfiltered v1 leaderboard contains 6,896 rows and the unfiltered v2 leaderboard contains 4,576 rows, so for obscure models in the tail the raw difference $\text{rank}_i^{v1} - \text{rank}_i^{v2}$ is dominated by the denominator shift rather than any genuine movement in relative position. On a per-rank basis, the unfiltered ANCOVA coefficient is roughly 50× smaller than the curated one (+0.00022 vs. +0.0101); on a per-percentile basis, the curated effect is instead $\sim 2.6\times$ larger. The curated sample is therefore the appropriate one for estimating the effect of movement among meaningful competitors, though the unfiltered replication still preserves the sign and significance of the relationship.

B.6 Per-horizon first-difference dynamics

Table 3 reports the first-difference coefficient horizon by horizon, using each individual post-shock snapshot as the outcome. The coefficient is positive at every horizon and generally increases once the measurement window lies fully post-shock, matching the event-study pattern in Figure 1.

Horizon	**Days post-shock**	N	$\hat{\beta}$	**SE**	p
Aug 08, 2024	42	116	+0.0060	0.0032	0.061
Aug 19, 2024	53	116	+0.0061	0.0032	0.054
Aug 27, 2024	61	116	+0.0072	0.0032	0.023
Sep 27, 2024	92	116	+0.0090	0.0033	0.007
Dec 27, 2024	183	116	+0.0074	0.0036	0.041
Mar 27, 2025	273	115	+0.0135	0.0036	< 0.001
Jun 27, 2025	365	113	+0.0137	0.0039	< 0.001

Table 3: First-difference regression of log downloads on ΔRank_i at each post-shock hub snapshot. Standard errors are HC3 robust. The July 24, 2024 snapshot (27 days post-shock) is excluded because its 30-day rolling window still overlaps the pre-shock period.

B.7 Interpretation

The identifying claim is local to the June 27, 2024 reranking. Because that rank change was induced by Hugging Face's evaluation-methodology choices rather than by anything the underlying models did, ΔRank_i is plausibly exogenous with respect to latent model quality conditional on pre-shock downloads. The absence of differential pre-shock trends over an eight-month pre-window (joint $F = 0.59$, $p = 0.56$) supports this interpretation. The post-shock coefficients should be read as reduced-form effects of exogenous public reranking on subsequent demand, allowing for endogenous propagation through adoption, tutorials, downstream references, and other ecosystem channels.