# The Power Game:
# A Discrete Dynamic Model of Compute and the AI Race

June 2026

**Abstract**

Artificial intelligence can look mature to consumers and still race at the frontier. The reason is that capability has two economic faces. On a fixed household menu, better reliability eventually saturates demand. For producers choosing ever longer tasks, the same reliability compounds value. This paper studies the resulting race in a discrete-time dynamic game. Each period, two laboratories choose training, compute prices clear against a saturated household-serving use, and the capability gap determines frontier payoffs. The finite-period game has Markov-perfect equilibria by backward induction; the discounted finite-state version has stationary Markov equilibria and is directly computable. The model preserves the paper's original mechanism while making the equilibrium object explicit: consumer saturation caps the serving value of compute, but producer stakes keep frontier training valuable, so scarce compute converts the race into input rents. Homogeneous Bertrand frontier services create a kink at parity, so the old smooth zero-sum threshold is not the right primitive; continuous-time HJB formulas are kept only as local approximations in the appendix.

**JEL:** L13, O31, O33, L40, D74.

## 1 Introduction

In many mature technology markets, racing slows once the marginal consumer stops rewarding further quality. Frontier artificial intelligence displays one consumer-facing symptom of that maturity–quality on many fixed tasks is saturating–and yet its firms continue to invest at the frontier. The puzzle is not that AI is already mature. It is that consumer saturation and frontier racing can coexist.

The paper's key concept is *task reliability.* A system with capability $q$ completes a single step of work correctly with some probability; a task of $n$ steps succeeds only if every step does, so success decays exponentially in task length. One primitive then has two faces.[1] Households

---

[1] The phrase echoes Cohen and Levinthal's (1989) "two faces of R&D" (Cohen and Levinthal, 1989); the dichotomy here is different–who controls the task frontier–but the cadence is borrowed knowingly.

use AI on a fixed menu of everyday tasks, so household quality saturates. Producers choose which tasks to attempt, so higher reliability lets them climb to longer and more valuable tasks. Proposition 1 gives the primitive: household value converges, while producer value grows as $A(q) = Ke^{aq}$.

The rest of the paper uses a discrete dynamic game. This is a modeling decision, not a cosmetic change. The continuous-time HJB model is elegant for local marginal formulas, but the payoff generated by homogeneous Bertrand frontier services is kinked at parity, and the dynamic game can have mixed or multiple equilibria. Those are not smoothness nuisances; they are part of the economics. The discrete model keeps the timing explicit: in each period, firms choose training, compute prices clear, the gap updates, and payoffs are realized. With finite states and finite action sets, finite-horizon games are solved by backward induction, and the infinite discounted game is a finite stochastic game with stationary Markov equilibria. Continuous time is therefore used only as a small-period approximation that explains local first-order conditions and comparative statics. It is not the foundation for the paper's global claims.

The contribution is to put an AI-specific supply side into a dynamic innovation race. The gap state, imitation force, and defended-lead logic are close to patent-race and step-by-step models (Harris and Vickers, 1985, 1987; Budd, Harris, and Vickers, 1993; Aghion et al., 2001; Horner, 2004; Acemoglu and Akcigit, 2012). The new object is the input market. Training is not bought at a fixed numeraire price; it bids against saturated household serving, connecting the race to input rents and exclusion logic (Krattenmaker and Salop, 1986; Chung, 1996; Goolsbee, 1998). The AI policy literature asks how upstream AI market power affects welfare and concentration (Korinek and Vipra, 2025; Athey and Scott Morton, 2025); this paper supplies a stripped-down mechanism for one part of that agenda.

The paper has five main formal statements. Proposition 1 derives the two faces. Proposition 2 builds the one-period and finite-period game. Proposition 3 gives compute-market discipline. Theorem 1 characterizes the finite-state discounted dynamic game and its computation. Proposition 4 records the composition/kink and access-pricing implications. Appendix A gives the derivations, including the continuous-time limit that connects the discrete model back to the HJB formulas.

## 2 Tasks, Reliability, and the Two Faces

**Environment.** Work is sequential: a task of length $n$ is a chain of $n$ steps, and one failed step kills it. A system of capability $q$ completes a step with probability $e^{-\lambda(q)}$, where $\lambda(q) = e^{-q}$. A task of length $n$ succeeds with probability $e^{-\lambda(q)n}$. Households face a fixed menu of tasks, represented by a finite value measure $v(dn)$ with total mass $\bar{C}$ and finite first moment

$\bar{N} = \int n\, v(dn)$. Household quality is

$$C(q) = \int e^{-\lambda(q)n}\, v(dn).$$

Producers choose the task length. A completed task of length $n$ is worth $B(n) = bn^a$, with $a > 0$. Producer value is

$$A(q) = \max_{n>0} bn^a e^{-\lambda(q)n}.$$

**Proposition 1** (Two faces of one primitive)**.** *(i)* Menu saturation. $C(q) \uparrow \bar{C}$ *and* $C'(q) \leq \bar{N}e^{-q} \to 0$*. The marginal household return to capability vanishes.*

*(ii)* Frontier compounding. *The producer's chosen task length is* $n^*(q) = ae^q$*, and*

$$A(q) = b(a/e)^a e^{aq} \equiv Ke^{aq}.$$

*Thus log producer value is linear in capability, and relative frontier value depends only on the capability gap.*

The menu face gives the saturated household spending flow $\bar{W} \equiv \bar{C}$. The frontier face gives the contest stakes. If the two laboratories have capabilities $q_1, q_2$, write the gap as $g = q_1 - q_2$. Then $A(q_1)/A(q_2) = e^{ag}$, so the gap alone determines relative frontier value.

## 3 A Small Discrete Game

This section starts with the smallest useful object: a one-period game. The finite-horizon and stationary models only repeat this block.

**Timing.** At the beginning of a period, the state is the capability gap $g$. Laboratories simultaneously choose training actions $x_i \geq 0$. Training uses compute $T_i = x_i^2/2$. The next gap is

$$g' = (1 - \mu)g + x_1 - x_2, \tag{1}$$

where $\mu \in (0, 1)$ is the one-period imitation rate. Frontier payoffs are earned on the post-training gap $g'$. This timing is important: training is costly today because it changes the capability gap on which today's frontier payoff and future value depend.

**Compute market.** Total training demand is $T = T_1 + T_2$. Household serving uses the residual compute $S = \bar{Q} - T$. In the constant-expenditure benchmark, saturated household spending $\bar{W}$ clears the rental market at

$$R(T) = \frac{\bar{W}}{\bar{Q} - T}, \qquad T < \bar{Q}. \tag{2}$$

The cold price is $r_0 = R(0) = \bar{W}/\bar{Q}$. A laboratory that owns compute faces the same opportunity cost: a chip used for training is a chip not rented to the serving market.

**Frontier payoff.** A share $s \in [0,1]$ of frontier value is one-shot. The probability that laboratory 1 solves first is logistic:

$$L(ag) = \frac{1}{1+e^{-ag}}.$$

The remaining share is repeatable and sold under homogeneous Bertrand competition: the leader earns the quality difference and the laggard earns zero. Laboratory 1's frontier payoff, net of the one-shot fair-split constant, is

$$\pi(g) = \omega \left[ s\left( L(ag) - \frac{1}{2} \right) + (1-s) 2 \sinh(ag/2)^{+} \right], \tag{3}$$

and laboratory 2's payoff is $\pi(-g)$. The one-period payoff is

$$u_1(g, x_1, x_2) = \pi(g') - R(T)\frac{x_1^2}{2}, \qquad u_2(g, x_1, x_2) = \pi(-g') - R(T)\frac{x_2^2}{2}. \tag{4}$$

**Proposition 2** (One-period and finite-period mechanics)**.** *Fix $\omega > 0$, $r_0 > 0$, and the payoff (3).*

1. *In the one-period game with continuous actions, inactive parity $(g, x_1, x_2) = (0,0,0)$ is not a Nash equilibrium. A small unilateral action $\varepsilon > 0$ gives laboratory 1 payoff*

$$\pi'_+(0)\varepsilon - \frac{r_0}{2}\varepsilon^2 + o(\varepsilon^2), \qquad \pi'_+(0) = \omega a(1 - 3s/4) > 0.$$

   *The benefit is first order and the compute cost is second order.*

2. *On a finite action grid, the same conclusion holds whenever the smallest positive action $\varepsilon$ is fine enough that $\pi(\varepsilon) > R(\varepsilon^2/2)\varepsilon^2/2$. If the grid is too coarse, inaction can survive as a discretization artifact.*

3. *For any finite horizon $H$, finite state grid $G$, finite action grid $X$, and bounded transition rule from (1) back to $G$, a Markov-perfect equilibrium exists and is computed by backward induction. At each date and state one solves a finite simultaneous-move normal-form game.*

The two- and three-period games add only continuation value. Let $V_{i,h}(g)$ be laboratory $i$'s value with $h$ periods remaining and $V_{i,0} \equiv 0$. With $h \geq 1$, the continuation game at state $g$ has payoffs

$$U_{i,h}(g, x_1, x_2) = u_i(g, x_1, x_2) + \beta V_{i,h-1}(g'), \tag{5}$$

where $g'$ is given by (1). A Markov-perfect equilibrium is obtained by choosing a Nash equilibrium of this finite matrix game at every state and date. In a two-period game, the one-period payoff slope is augmented by the slope of $V_{i,1}$. In a three-period game, the same recursion repeats. This is the mechanical reason the finite model is easier to understand: all dynamic incentives are continuation values in finite normal-form games.

## 4 Compute-Market Discipline

**Proposition 3** (Compute-market discipline)**.** *For every period and every action profile with $T < \bar{Q}$:*

1. *The rental price is $R(T) = \bar{W}/(\bar{Q} - T)$. Physical training is bounded by capacity.*
2. *Training expenditure is*

$$E_T = R(T)T = \frac{\bar{W}T}{\bar{Q} - T}.$$

   *Thus unbounded training expenditure or rental rates require $T \uparrow \bar{Q}$; escalation is paid as compute scarcity rather than unbounded physical chip quantities.*

3. *If a continuation value is differentiable and the period is locally approximated by a marginal action, the marginal training benefit is the marginal value of the induced gap increase, while the marginal cost is $r_0 x + o(x)$. This recovers the first-order-benefit, second-order-cost logic of Proposition 2.*

The result is accounting, not an equilibrium selection theorem. It says where the money goes: once training demand presses on capacity, higher stakes show up as higher compute rents.

## 5 The Dynamic Discrete Game

The infinite-horizon version is a discounted stochastic game on finite grids. Let $G$ be a symmetric finite gap grid and $X$ a finite action grid containing zero. A transition kernel $P(g'' \mid g, x_1, x_2)$ maps the post-training gap $g' = (1 - \mu)g + x_1 - x_2$ back to $G$ by projection or adjacent interpolation. Payoffs are (4). A stationary Markov strategy maps each $g \in G$ to a mixed action over $X$.

**Theorem 1** (Finite-state dynamic equilibrium and computation)**.** *In the finite-state, finite-action discounted game with $\beta \in (0, 1)$, bounded payoffs, and transition kernel $P$, a stationary mixed Markov-perfect equilibrium exists. It is characterized by value functions $V_1, V_2$ such that, for each state $g$, the mixed actions $(\sigma_1(g), \sigma_2(g))$ are a Nash equilibrium of the finite continuation game with payoff matrices*

$$A_i(g; x_1, x_2) = u_i(g, x_1, x_2) + \beta \sum_{g'' \in G} P(g'' \mid g, x_1, x_2) V_i(g'').$$

*The equilibrium can be computed on a chosen grid by solving these finite continuation games inside a value-iteration, policy-iteration, or homotopy routine. For any finite horizon, the same matrices give an exact backward-induction computation.*

The theorem is the main reason to put discrete time in the body. The equilibrium object is no longer a local differentiable HJB branch; it is a Markov-perfect equilibrium of a finite dynamic game. Multiplicity does not disappear. A finite simultaneous-move game can have several equilibria at a state, and a dynamic game can inherit multiple Markov equilibria. But that is an economic selection problem, not an analytic regularity problem hidden inside continuous time. In this sense the discrete model is more conservative: it gives up closed-form global thresholds in exchange for an equilibrium object whose existence and computation do not require differentiability at parity.

The verification package computes a three-period version of this game on a 41-point gap grid and a 7-point action grid. The calculation is not a calibration. It is a transparent check that the finite-period object can be solved exactly by backward induction once a grid and an equilibrium-selection rule are chosen.

## 6 Composition, Escalation, and Access

The discrete model keeps the same economic composition logic as the continuous model. The issue is whether frontier value is solved once or sold repeatedly.

**Proposition 4** (Composition and market edges)**.** *Let* $L(z) = 1/(1+e^{-z})$*.*

1. *If one-shot hazards are proportional to* $A(q_i)^\theta$*, the one-shot payoff per unit of frontier value is* $L(\theta a g) - 1/2$*, with expansion*

$$L(\theta a g) - 1/2 = \frac{\theta a}{4} g - \frac{(\theta a)^3}{48} g^3 + O(g^5).$$

2. *The homogeneous-Bertrand repeatable payoff for laboratory 1 is* $2\sinh(ag/2)^+$*. Hence, for* $s < 1$*, the full payoff is kinked at parity:*

$$\pi'_-(0) = \omega \frac{s\theta a}{4}, \qquad \pi'_+(0) = \omega a \left(1 - s + \frac{s\theta}{4}\right).$$

   *The old smooth* $s^* = 1/2$ *threshold is therefore not a microfounded Bertrand result; it belongs to a different odd payoff projection.*

3. *In any equilibrium action profile,* $T < \bar{Q}$*. If a sequence of dynamic equilibria has training demand* $T \to \bar{Q}$*, then* $R(T) \to \infty$ *and* $E_T \to \infty$*. Thus expenditure escalation is absorbed by compute-owner rents.*

4. *In the static no-learning access benchmark, a leader with quality $q_L$ selling to a follower with quality $q_F$ charges the raw-output limit price*

$$P_0(g) = \frac{A(q_L)}{A(q_F)} = e^{ag}, \qquad g = q_L - q_F.$$

*The raw-unit markup is $e^{ag} - 1$ and the ad-valorem revenue margin is $1 - e^{-ag}$. If access also raises buyer capability by leakage $\ell(E)$, and the seller cannot charge for that learning benefit, selling is privately profitable only if*

$$E(1 - e^{-ag}) \geq \ell(E)\Delta V(g),$$

*where $\Delta V(g)$ is the seller's dynamic value loss from the buyer's extra capability.*

The access result is intentionally separated from the dynamic equilibrium computation. It gives a pricing implication conditional on the value difference created by the dynamic game. Near parity, if the leader's dynamic loss from leakage is positive, access is refused; at a large enough lead, access can resume if the marginal value of further separation is small.

## 7 Implications and Required Evidence

The model supplies mathematical implications, not empirical estimates. First, access prices identify the value elasticity only under no-learning or separately priced learning:

$$\log P_0 = a \log(H_L/H_F),$$

where $H_L/H_F$ is the ratio of reliable task horizons. Estimating $a$ requires access prices, task-horizon measures, and corrections for missing prices when access is refused.

Second, the compute-market mechanism predicts that training demand is visible in input markets. The relevant evidence is training-class compute allocation, supplier margins, capacity pre-booking, and rental prices for frontier compute. Low consumer serving margins do not imply peace; they determine the opportunity cost of training.

Third, the discrete model clarifies what remains unresolved. Equilibrium existence is easy on finite grids, and finite horizons are mechanically computable, but equilibrium selection is real. Symmetric noise, public correlation devices, or refinement assumptions are economic additions, not algebraic details. Continuous-time formulas are useful summaries of local forces, but the object to estimate or simulate is the Bellman-Nash system on a specified state and action grid. A welfare version also needs extra primitives: spillovers from frontier advances, duplicated-training costs, whether compute-owner rents are transfers or distortions, and whether access leakage is socially valuable or harmful.

# A Step-by-Step Derivations

## A.1 Two Faces

Start with the household object

$$C(q) = \int e^{-\lambda(q)n} v(dn), \qquad \lambda(q) = e^{-q}.$$

Step 1: as $q \to \infty$, $e^{-q} \to 0$, so $\lambda(q) \to 0$. For each fixed $n$,

$$e^{-\lambda(q)n} \to e^0 = 1.$$

Step 2: because $\lambda(q) \downarrow 0$, the integrand $e^{-\lambda(q)n}$ increases monotonically to one. Since $v$ is a finite measure, monotone convergence gives

$$\lim_{q\to\infty} C(q) = \int 1\, v(dn) = \bar{C}.$$

Step 3: differentiate the integrand. Since $\lambda'(q) = -e^{-q} = -\lambda(q)$,

$$\frac{d}{dq} e^{-\lambda(q)n} = e^{-\lambda(q)n}\{-n\lambda'(q)\} = ne^{-q}e^{-\lambda(q)n}.$$

Step 4: integrate this derivative:

$$C'(q) = e^{-q} \int ne^{-\lambda(q)n} v(dn).$$

Step 5: use $e^{-\lambda(q)n} \leq 1$ and $\bar{N} = \int n\, v(dn)$:

$$0 \leq C'(q) \leq e^{-q}\bar{N} \to 0.$$

This proves household saturation and vanishing marginal household value.

For producers, solve

$$A(q) = \max_{n>0} bn^a e^{-\lambda(q)n}.$$

Step 1: take logs of the objective, dropping the constant $\log b$:

$$\log b + a \log n - \lambda(q)n.$$

Step 2: differentiate with respect to $n$:

$$\frac{a}{n} - \lambda(q).$$

Step 3: set the derivative equal to zero:

$$\frac{a}{n} = \lambda(q) \quad \Longrightarrow \quad n^*(q) = \frac{a}{\lambda(q)}.$$

Step 4: substitute $\lambda(q) = e^{-q}$:

$$n^*(q) = ae^q.$$

Step 5: the second derivative of the log objective is

$$-\frac{a}{n^2} < 0,$$

so the first-order condition gives the unique maximum. Step 6: substitute $n^*(q) = a/\lambda(q)$ into the objective:

$$A(q) = b\left(\frac{a}{\lambda(q)}\right)^a \exp\left[-\lambda(q)\frac{a}{\lambda(q)}\right].$$

Step 7: simplify the exponential term:

$$\exp\left[-\lambda(q)\frac{a}{\lambda(q)}\right] = e^{-a}.$$

Step 8: substitute $\lambda(q) = e^{-q}$:

$$\left(\frac{a}{\lambda(q)}\right)^a = a^a e^{aq}.$$

Therefore

$$A(q) = b(a/e)^a e^{aq} \equiv Ke^{aq}.$$

For two laboratories,

$$\frac{A(q_1)}{A(q_2)} = \frac{Ke^{aq_1}}{Ke^{aq_2}} = e^{a(q_1-q_2)} = e^{ag}.$$

## A.2 One-Period Payoff and No-Inactive-Settlement

At state $g$, the next gap is

$$g' = (1-\mu)g + x_1 - x_2.$$

Training uses $T_i = x_i^2/2$, so total training demand is

$$T = \frac{x_1^2 + x_2^2}{2}.$$

Residual serving compute is $S = \bar{Q} - T$. Constant household spending $\bar{W}$ implies rental price

$$R(T) = \frac{\bar{W}}{S} = \frac{\bar{W}}{\bar{Q} - T}.$$

Laboratory 1's one-period payoff is

$$u_1(g, x_1, x_2) = \pi(g') - R(T)\frac{x_1^2}{2}.$$

At parity, set $g = 0$. If both firms are inactive, $x_1 = x_2 = 0$, then

$$g' = 0, \qquad T = 0, \qquad u_1(0, 0, 0) = \pi(0).$$

The payoff normalization gives $\pi(0) = 0$, so

$$u_1(0, 0, 0) = 0.$$

Now let laboratory 1 deviate to $x_1 = \varepsilon > 0$, while laboratory 2 remains at $x_2 = 0$. Then

$$g' = (1 - \mu)0 + \varepsilon - 0 = \varepsilon,$$

and

$$T = \frac{\varepsilon^2}{2}.$$

The deviating payoff is

$$u_1(0, \varepsilon, 0) = \pi(\varepsilon) - R(\varepsilon^2/2)\frac{\varepsilon^2}{2}.$$

Therefore the payoff gain from deviation is

$$u_1(0, \varepsilon, 0) - u_1(0, 0, 0) = \pi(\varepsilon) - R(\varepsilon^2/2)\frac{\varepsilon^2}{2}.$$

Because $R$ is continuous at zero training,

$$R(\varepsilon^2/2) = R(0) + O(\varepsilon^2) = r_0 + O(\varepsilon^2).$$

Multiplying by $\varepsilon^2/2$ gives

$$R(\varepsilon^2/2)\frac{\varepsilon^2}{2} = \frac{r_0}{2}\varepsilon^2 + O(\varepsilon^4).$$

For small positive $g$, the payoff expansion is

$$\pi(g) = \pi'_+(0)g + o(g).$$

Substituting $g = \varepsilon$ gives

$$\pi(\varepsilon) = \pi'_+(0)\varepsilon + o(\varepsilon).$$

Thus

$$u_1(0, \varepsilon, 0) - u_1(0, 0, 0) = \pi'_+(0)\varepsilon - \frac{r_0}{2}\varepsilon^2 + o(\varepsilon).$$

The baseline payoff has

$$\pi'_+(0) = \omega a(1 - 3s/4).$$

Since $s \in [0, 1]$,

$$1 - \frac{3s}{4} \geq \frac{1}{4} > 0,$$

so $\pi'_+(0) > 0$ whenever $\omega > 0$. The gain is positive for all sufficiently small $\varepsilon > 0$. Hence zero training at parity is not a one-period Nash equilibrium in the continuous-action game.

On a finite grid, the smallest positive action is some $\varepsilon$. The same deviation works if and only if

$$\pi(\varepsilon) > R(\varepsilon^2/2)\frac{\varepsilon^2}{2}.$$

This is why a coarse grid can create artificial inaction: the first available $\varepsilon$ may be too large for the first-order approximation to dominate the quadratic cost.

## A.3 Finite-Horizon Backward Induction

Let $H$ be the horizon. Let $V_{i,h}(g)$ denote laboratory $i$'s value when there are $h$ periods remaining. The terminal condition is

$$V_{i,0}(g) = 0 \quad \text{for all } g.$$

Suppose $V_{i,h-1}$ has already been computed. At state $g$, each pair of current actions $(x_1, x_2)$ implies

$$g' = (1 - \mu)g + x_1 - x_2.$$

If $g'$ is not exactly on the finite grid, the computational model maps it back to the grid by projection or interpolation. Denote the resulting continuation value by $V_{i,h-1}(g')$. The total current-plus-continuation payoff is

$$U_{i,h}(g, x_1, x_2) = u_i(g, x_1, x_2) + \beta V_{i,h-1}(g').$$

For a fixed state $g$, the action sets are finite, so the matrices

$$\left[U_{1,h}(g, x_1, x_2)\right]_{x_1,x_2}, \qquad \left[U_{2,h}(g, x_1, x_2)\right]_{x_1,x_2}$$

define a finite normal-form game. A mixed Nash equilibrium exists in every finite normal-form game. Choose one equilibrium at each state; call it

$$\sigma_{1,h}(g), \sigma_{2,h}(g).$$

The value is the expected payoff under that equilibrium:

$$V_{i,h}(g) = \mathbb{E}_{\sigma_{1,h}(g),\sigma_{2,h}(g)}\left[U_{i,h}(g, x_1, x_2)\right].$$

This completes the induction step. Starting from $V_{i,0} = 0$, applying this step for $h = 1, 2, \ldots, H$ constructs a finite-horizon Markov-perfect equilibrium.

For $H = 1$, the continuation term is zero:

$$U_{i,1}(g, x_1, x_2) = u_i(g, x_1, x_2).$$

For $H = 2$,

$$U_{i,2}(g, x_1, x_2) = u_i(g, x_1, x_2) + \beta V_{i,1}(g').$$

For $H = 3$,

$$U_{i,3}(g, x_1, x_2) = u_i(g, x_1, x_2) + \beta V_{i,2}(g').$$

Thus the two- and three-period games introduce no new mathematical object; they repeatedly add one more continuation-value term.

### A.4 Stationary Finite Game

Now take a finite state grid $G$, a finite action grid $X$, and a discount factor $\beta \in (0, 1)$. Let $P(g'' \mid g, x_1, x_2)$ be the transition probability from current state $g$ and actions $(x_1, x_2)$ to next grid state $g''$. A stationary Markov strategy maps each state $g$ to a mixed action over $X$.

Given candidate value functions $V_1, V_2$, define the continuation payoff matrices

$$A_i(g; x_1, x_2) = u_i(g, x_1, x_2) + \beta \sum_{g'' \in G} P(g'' \mid g, x_1, x_2) V_i(g'').$$

At a stationary Markov equilibrium, the mixed actions $(\sigma_1(g), \sigma_2(g))$ must be a Nash equilibrium of this matrix game at every $g$. The associated values must satisfy

$$V_i(g) = \mathbb{E}_{\sigma_1(g),\sigma_2(g)}\left[A_i(g; x_1, x_2)\right].$$

These two conditions are exactly the Bellman-Nash fixed point. Since the state set is finite, the action sets are finite, one-period payoffs are bounded, and $\beta < 1$, Fink's theorem for discounted stochastic games gives existence of a stationary mixed equilibrium.

### A.5 Compute-Market Discipline

Total training demand is

$$T = T_1 + T_2 = \frac{x_1^2 + x_2^2}{2}.$$

Residual serving compute is

$$S = \bar{Q} - T.$$

The constant-expenditure household side spends $\bar{W}$, so the rental price per unit of compute is

$$R(T) = \frac{\bar{W}}{S} = \frac{\bar{W}}{\bar{Q} - T}.$$

Training expenditure is price times quantity:

$$E_T = R(T)T.$$

Substitute the rental price:

$$E_T = \frac{\bar{W}}{\bar{Q} - T} T = \frac{\bar{W}T}{\bar{Q} - T}.$$

If $T < \bar{Q}$, physical training is bounded by capacity. If $T$ stays below capacity by some $\delta > 0$, so $T \le \bar{Q} - \delta$, then

$$R(T) \le \frac{\bar{W}}{\delta},$$

and

$$E_T \le \frac{\bar{W}(\bar{Q} - \delta)}{\delta}.$$

Therefore $R(T)$ or $E_T$ can diverge only if

$$T \uparrow \bar{Q}.$$

This proves the scarcity-rent interpretation: high stakes cannot create more than $\bar{Q}$ physical compute in the model; they can only bid up its price.

## A.6 Composition and Kink

The one-shot probability is

$$L(\theta a g) = \frac{1}{1 + e^{-\theta a g}}.$$

The logistic derivatives at zero are

$$L(0) = \frac{1}{2}, \qquad L'(0) = \frac{1}{4}, \qquad L''(0) = 0, \qquad L'''(0) = -\frac{1}{8}.$$

Taylor expansion around zero gives

$$L(z) = \frac{1}{2} + \frac{1}{4}z - \frac{1}{48}z^3 + O(z^5).$$

Set $z = \theta a g$:

$$L(\theta a g) - \frac{1}{2} = \frac{\theta a}{4} g - \frac{(\theta a)^3}{48} g^3 + O(g^5).$$

The repeatable Bertrand term is

$$2\sinh(ag/2)^+.$$

If $g < 0$, then $\sinh(ag/2) < 0$, so

$$2\sinh(ag/2)^+ = 0.$$

If $g > 0$, then

$$2\sinh(ag/2)^+ = 2\sinh(ag/2).$$

Using

$$\sinh z = z + \frac{z^3}{6} + O(z^5)$$

with $z = ag/2$, we get

$$2\sinh(ag/2) = 2\left(\frac{ag}{2} + \frac{(ag/2)^3}{6} + O(g^5)\right) = ag + \frac{a^3 g^3}{24} + O(g^5).$$

Therefore the left derivative at parity comes only from the one-shot part:

$$\pi'_-(0) = \omega \frac{s\theta a}{4}.$$

The right derivative includes both one-shot and repeatable terms:

$$\pi'_+(0) = \omega \left[ s\frac{\theta a}{4} + (1-s)a \right] = \omega a \left(1 - s + \frac{s\theta}{4}\right).$$

If $s < 1$, then

$$\pi'_+(0) - \pi'_-(0) = \omega a (1-s) > 0.$$

Thus the homogeneous-Bertrand repeatable payoff is kinked at parity.

### A.7 Access Pricing

Let the leader have capability $q_L$ and the follower $q_F$, with gap

$$g = q_L - q_F > 0.$$

Producer value is $A(q) = Ke^{aq}$. Therefore the quality ratio is

$$\frac{A(q_L)}{A(q_F)} = \frac{Ke^{aq_L}}{Ke^{aq_F}} = e^{a(q_L - q_F)} = e^{ag}.$$

If the leader charges raw-unit access price $P$, the buyer gets effective quality $A(q_L)/P$ per dollar. Self-supply gives effective quality $A(q_F)$ per dollar. Static no-learning indifference requires

$$\frac{A(q_L)}{P_0} = A(q_F).$$

Solving for $P_0$,

$$P_0 = \frac{A(q_L)}{A(q_F)} = e^{ag}.$$

The raw-unit markup over normalized cost one is

$$P_0 - 1 = e^{ag} - 1.$$

The ad-valorem revenue margin is

$$\frac{P_0 - 1}{P_0} = 1 - \frac{1}{P_0} = 1 - e^{-ag}.$$

Gross expenditure $E$ buys raw quantity $E/P_0$. With normalized unit cost one, the seller's cost is $E/P_0$, revenue is $E$, and static rent is

$$E - \frac{E}{P_0} = E(1 - e^{-ag}).$$

If access creates buyer learning $\ell(E)$, and the seller's dynamic value loss per unit of that learning is $\Delta V(g)$, then the dynamic loss is

$$\ell(E)\Delta V(g).$$

Selling is privately profitable exactly when static rent weakly exceeds dynamic loss:

$$E(1 - e^{-ag}) \geq \ell(E)\Delta V(g).$$

## A.8 Continuous-Time Approximation

Let $\Delta$ be the period length and set

$$\beta = e^{-\rho\Delta} = 1 - \rho\Delta + O(\Delta^2).$$

Scale the gap law as

$$g' = g + \Delta(x_1 - x_2 - \mu g).$$

Scale current payoff as

$$\Delta\left[\pi(g) - r\frac{x_1^2}{2}\right].$$

The discrete Bellman equation for laboratory 1 is

$$V(g) = \max_{x_1 \geq 0} \left\{ \Delta \left[ \pi(g) - r\frac{x_1^2}{2} \right] + \beta V(g + \Delta(x_1 - x_2 - \mu g)) \right\}.$$

Use the Taylor expansion

$$V(g + \Delta d) = V(g) + \Delta d V'(g) + O(\Delta^2),$$

where $d = x_1 - x_2 - \mu g$. Substitute this and $\beta = 1 - \rho\Delta + O(\Delta^2)$:

$$V(g) = \max_{x_1 \geq 0} \left\{ \Delta \left[ \pi(g) - r\frac{x_1^2}{2} \right] + (1 - \rho\Delta) \left[ V(g) + \Delta d V'(g) \right] + O(\Delta^2) \right\}.$$

Expand the right side:

$$V(g) = \max_{x_1 \geq 0} \left\{ V(g) + \Delta \left[ \pi(g) - r\frac{x_1^2}{2} + dV'(g) - \rho V(g) \right] + O(\Delta^2) \right\}.$$

Subtract $V(g)$, divide by $\Delta$, and let $\Delta \to 0$:

$$0 = \max_{x_1 \geq 0} \left\{ \pi(g) - r\frac{x_1^2}{2} + (x_1 - x_2 - \mu g)V'(g) - \rho V(g) \right\}.$$

Rearranging gives

$$\rho V(g) = \pi(g) + \max_{x_1 \geq 0} \left\{ x_1 V'(g) - r\frac{x_1^2}{2} \right\} - x_2 V'(g) - \mu g V'(g).$$

In a symmetric Markov profile, $x_2 = x(-g)$, so

$$\rho V(g) = \pi(g) + \max_{x \geq 0} \left\{ x V'(g) - r\frac{x^2}{2} \right\} - x(-g)V'(g) - \mu g V'(g).$$

The maximand in $x$ is

$$xV'(g) - r\frac{x^2}{2}.$$

Its derivative is

$$V'(g) - rx.$$

If $V'(g) > 0$, the interior maximizer is

$$x = \frac{V'(g)}{r}.$$

If $V'(g) \leq 0$, the nonnegativity constraint binds and $x = 0$. Therefore

$$x = \frac{[V'(g)]^+}{r}.$$

This derivation shows exactly what continuous time buys: the clean local marginal formula. The equilibrium characterization in the main text does not require it.

## References

Harris, C., and J. Vickers (1985): "Perfect Equilibrium in a Model of a Race," *Review of Economic Studies*, 52(2), 193–209.

Harris, C., and J. Vickers (1987): "Racing with Uncertainty," *Review of Economic Studies*, 54(1), 1–21.

Budd, C., C. Harris, and J. Vickers (1993): "A Model of the Evolution of Duopoly: Does the Asymmetry between Firms Tend to Increase or Decrease?" *Review of Economic Studies*, 60(3), 543–573.

Horner, J. (2004): "A Perpetual Race to Stay Ahead," *Review of Economic Studies*, 71(4), 1065–1088.

Aghion, P., C. Harris, P. Howitt, and J. Vickers (2001): "Competition, Imitation and Growth with Step-by-Step Innovation," *Review of Economic Studies*, 68(3), 467–492.

Konrad, K. A., and D. Kovenock (2009): "Multi-Battle Contests," *Games and Economic Behavior*, 66(1), 256–274.

Krattenmaker, T. G., and S. C. Salop (1986): "Anticompetitive Exclusion: Raising Rivals' Costs to Achieve Power over Price," *Yale Law Journal*, 96(2), 209–293.

Denicolò, V., and L. A. Franzoni (2010): "On the Winner-Take-All Principle in Innovation Races," *JEEA*, 8(5), 1133–1158.

Fink, A. M. (1964): "Equilibrium in a Stochastic n-Person Game," *Journal of Science of the Hiroshima University, Series A-I*, 28(1), 89–93.

Korinek, A., and J. Vipra (2025): "Concentrating Intelligence: Scaling and Market Structure in Artificial Intelligence," *Economic Policy*, 40(121), 225–256, doi:10.1093/epolic/eiae057.

Athey, S., and F. Scott Morton (2025): "Artificial Intelligence, Competition, and Welfare," NBER Working Paper No. 34444, November, doi:10.3386/w34444.

Acemoglu, D., and U. Akcigit (2012): "Intellectual Property Rights Policy, Competition and Innovation," *JEEA*, 10(1), 1–42.

Chung, T.-Y. (1996): "Rent-Seeking Contest When the Prize Increases with Aggregate Efforts," *Public Choice*, 87(1–2), 55–66.

Goolsbee, A. (1998): “Does Government R&D Policy Mainly Benefit Scientists and Engineers?” *American Economic Review*, 88(2), 298–302.

Cohen, W. M., and D. A. Levinthal (1989): “Innovation and Learning: The Two Faces of R&D,” *Economic Journal*, 99(397), 569–596.

Goettler, R., and B. Gordon (2011): “Does AMD Spur Intel to Innovate More?” *Journal of Political Economy*, 119(6), 1141–1200.

Kwa, T., et al. (METR) (2025): “Measuring AI Ability to Complete Long Software Tasks,” *NeurIPS 2025*, arXiv:2503.14499.

METR (2026): “Task-Completion Time Horizons of Frontier AI Models,” public time-horizon series, updated May 8, 2026, https://metr.org/time-horizons/.