Release Dynamics with Private Breakthroughs

A minimal timing game for strategic AGI-style releases

Apoorva Lal

2026-07-12

Table of contents

1 One useful workhorse

A sensible first model is a **discrete-time stochastic preemption game with private breakthrough opportunities**. It is a patent-race or real-options timing game relabelled as a release race:

- each firm privately receives opportunities to release;
- the public only sees `act` or `wait`;
- the most recent actor is the current leader and captures a per-period surplus;
- simultaneous releases dissipate value, so firms dislike co-locating.

This is deliberately less literal than a full latent-threshold model. The baseline treats a threshold crossing as a one-period private action opportunity. That choice makes the model small enough to solve and to inspect. The more literal "private button persists until pressed" version is a belief-state extension; it is the right next model if the data or application really requires strategic stockpiling of unreleased ideas.

The main message is that the verbal description is **not yet enough to solve a unique game**. It gives the action labels and some payoff intuition, but not the primitives that pin down values, beliefs, state transitions, or the treatment of ties. Once those are supplied, the problem is a standard discounted dynamic Bayesian game.

2 What is underspecified?

The following primitives matter. Changing any of them can change the equilibrium.

1. **Horizon and discounting.** Is the game finite, infinite discounted, or average-payoff? Is there a terminal prize?
2. **Surplus process.** Is the per-period surplus constant s, deterministic s_t, stochastic, or affected by the identity of the leader?
3. **Leadership transition.** If one firm acts, does it immediately become the leader for the current period, the next period, or both?
4. **Tie resolution.** If both firms act, is the current surplus split as $s/2 - k$ only for that period? Who leads next period? Is the tie state persistent?
5. **Action feasibility.** Can a firm act only after a private idea arrival? Can the current leader act defensively? Does action consume the idea?
6. **Arrival process.** Are breakthroughs iid Bernoulli arrivals, Poisson arrivals, Markov thresholds, correlated across firms, or strategically influenced by effort?
7. **Persistence of private readiness.** Once the private `act` button appears, does it persist until used, expire after one period, or decay?
8. **Information.** Does `wait` mean "no idea" or "strategic withholding"? Are simultaneous releases observed as simultaneous? Are failed/aborted releases observed?
9. **Equilibrium concept.** A one-shot simultaneous-move model, a perfect Bayesian equilibrium, a Markov perfect Bayesian equilibrium, and an equilibrium with public beliefs are not the same object.

10. **Identification from data.** An observed wait is censored: it may be caused by no breakthrough or by a firm with a button choosing not to press it. Without restrictions, action histories alone cannot separately identify arrival hazards and strategic withholding.

The model below chooses the missing primitives explicitly. It should be read as a clean starting point, not as the only possible formalization.

3 A defensible closure: ten assumptions

Here is the baseline closure I would actually defend. The goal is not to pretend that these assumptions are innocuous. The goal is to make one mathematically clean object whose moving parts are visible, whose equilibrium is computable, and whose failure modes are easy to name.

1. **Infinite discounted time, no terminal prize.** Use an infinite horizon with common $\delta \in (0,1)$ and no exogenous deadline. This is the conservative stationary benchmark: it avoids finite-horizon end effects and isolates the strategic timing force. A deadline, launch window, or terminal patent prize can be added later as a finite-horizon perturbation.
2. **Constant common surplus.** Set the flow surplus to a constant $s > 0$, identical across firms and calendar time. This makes "being in the lead" the object of competition rather than adding a second state variable for demand or attention. Empirically, this is the right first normalization when the available data are action times rather than a high-frequency market-size process.
3. **Immediate leadership transfer.** A unilateral action captures the current period's surplus and makes the actor the next-period leader. This is the cleanest convention for a discrete-time model: acting is both the public release and the beginning of exploitation. If implementation lag matters, shift the surplus to $t+1$; the Bellman equations change mechanically.
4. **Collisions are one-period events with random continuation.** If both act, each receives $\tau = s/2 - k$ for that period and next period's leader is selected by a fair public lottery. This separates the collision payoff from the continuation allocation. It is deliberately neutral: neither player is favored after a tie, and there is no hidden persistent co-leader state.
5. **Acting is feasible only with a fresh private opportunity.** A firm can press 'act' only if its private threshold has crossed in that period. The current leader can also receive such an opportunity and act defensively. Action consumes the opportunity. This makes both offensive release and defensive co-release possible without giving firms arbitrary release power.

6. **Breakthroughs are iid Bernoulli arrivals.** In the symmetric baseline, $R_{it} \sim$ Bernoulli(λ) independently across firms and periods. This is the discrete-time analogue of independent Poisson breakthroughs. It is not a claim that ideas are literally iid; it is the minimal hazard model against which correlation, learning, and effort choice should be measured.

7. **Private buttons expire after one period.** A threshold crossing generates a one-period release opportunity. If the firm waits, the opportunity disappears. This is the major tractability assumption: it prevents private inventories of unreleased ideas from becoming a hidden state. The persistent-button model is treated separately below because it is a belief-state game, not a three-value fixed point.

8. **Only public actions are observed.** Everyone observes `act' and`wait', including whether both firms acted in the same period. A wait does not reveal whether the firm lacked an opportunity or strategically withheld one. In the one-period baseline this ambiguity does not create a public belief state because opportunities expire; in the persistent-button extension it does.

9. **Stationary symmetric Markov perfect Bayesian equilibrium.** Strategies depend only on the public leader state and the player's current private opportunity. This is the natural equilibrium concept for the infinite discounted baseline with recurrent private types. Symmetry is an expositional restriction, not a conceptual one; asymmetric hazards λ_i or payoffs simply replace three values with player-specific values.

10. **No structural identification from actions alone.** Treat the model as a disciplined theory object unless extra variation is available. Binary action histories alone cannot distinguish low arrival rates from strategic waiting, high collision costs, or pessimistic beliefs about the opponent's readiness. This is a modeling assumption about what the exercise can and cannot deliver: it characterizes counterfactuals conditional on primitives; it does not magically recover those primitives.

These ten assumptions turn the verbal story into a standard discounted dynamic Bayesian game. The restrictions are transparent enough that one can relax them one at a time: add a terminal date, make s_t stochastic, let opportunities persist, add correlated arrivals, or estimate beliefs from auxiliary signals. The baseline earns its keep by showing the strategic logic before those extensions are layered on.

4 Baseline model

Time is discrete, $t = 0, 1, 2, ...$ There are two firms $i \in \{1, 2\}$. The public state at the beginning of period t is

$$\ell_t \in \{0, 1, 2\},$$

where $\ell_t = 0$ means there is no current leader and $\ell_t = i$ means firm i is the current leader.

Each period, before actions, firm i privately observes a release opportunity

$$R_{it} \in \{0,1\}, \qquad \Pr(R_{it} = 1) = \lambda_i.$$

The baseline assumes iid arrivals across firms and periods. A firm can act only when $R_{it} = 1$. A firm without an opportunity must wait. An opportunity lasts for one period in the baseline.

The action is

$$a_{it} \in \{0,1\}, \qquad a_{it} \le R_{it},$$

where $a_{it} = 1$ means `act' and $a_{it}=0$ means`wait'. The action pair is publicly observed.

Let $s > 0$ be the per-period surplus. If exactly one firm acts, that firm captures the current period's surplus and becomes the leader:

$$(a_{it}, a_{jt}) = (1,0) \quad \Rightarrow \quad u_{it} = s, \quad u_{jt} = 0, \quad \ell_{t+1} = i.$$

If nobody acts, the current leader, if any, keeps the period's surplus:

$$(a_{it}, a_{jt}) = (0,0),\ \ell_t = i \quad \Rightarrow \quad u_{it} = s, \quad u_{jt} = 0, \quad \ell_{t+1} = i.$$

If there is no current leader and nobody acts, both payoffs are zero and $\ell_{t+1} = 0$.

If both act, the period is a collision. Each firm receives

$$\tau = \frac{s}{2} - k,$$

where $k \ge 0$ is a co-location haircut. For continuation, the baseline breaks the tie uniformly:

$$\Pr(\ell_{t+1} = 1 \mid a_{1t} = a_{2t} = 1) = \Pr(\ell_{t+1} = 2 \mid a_{1t} = a_{2t} = 1) = \frac{1}{2}.$$

The firms maximize expected discounted payoff with discount factor $\delta \in (0,1)$.

This tie rule is only a normalization. If simultaneous release creates a persistent co-leader state, a random market winner, or a delayed resolution, that should be added as an explicit public state.

5 Equilibrium concept

The natural concept is a **stationary Markov perfect Bayesian equilibrium**. Strategies condition on the public leader state and the player's private opportunity:

$$\sigma_i(a_i = 1 \mid \ell, R_i).$$

Feasibility requires $\sigma_i(1 \mid \ell, 0) = 0$. In the iid one-period-arrival baseline, there is no persistent hidden state to filter. Beliefs over the opponent's current opportunity are always the common prior λ_j.

For a symmetric stationary equilibrium, write:

- x: probability that the current leader acts when it has an opportunity;
- y: probability that the follower acts when it has an opportunity;
- z: probability that a firm acts when there is no current leader and it has an opportunity.

Let V_L be the value of being current leader at the start of a period, V_F the value of being follower, and V_0 the value for either firm when there is no current leader. Let

$$M = \frac{V_L + V_F}{2}.$$

Given (x, y, z), the leader and follower values solve:

$$\begin{aligned}
V_L &= (1-\lambda y)(s+\delta V_L) + (1-\lambda x)\lambda y\,\delta V_F + \lambda x\lambda y(\tau+\delta M),\\
V_F &= \lambda y(1-\lambda x)(s+\delta V_L) + (1-\lambda y)\delta V_F + \lambda x\lambda y(\tau+\delta M),\\
V_0 &= (\lambda z)(1-\lambda z)(s+\delta V_L) + (1-\lambda z)(\lambda z)\delta V_F\\
&\quad + (\lambda z)^2(\tau+\delta M) + (1-\lambda z)^2\delta V_0.
\end{aligned}$$

The one-period deviation inequalities characterize equilibrium. Define the conditional gains from pressing the button rather than waiting:

$$\begin{aligned}
\Delta_L &= \lambda y\left[\tau + \delta M - \delta V_F\right],\\
\Delta_F &= (1-\lambda x)(s+\delta V_L) + \lambda x(\tau+\delta M) - \delta V_F,\\
\Delta_0 &= (1-\lambda z)(s+\delta V_L) + \lambda z(\tau+\delta M)\\
&\quad - \left[(1-\lambda z)\delta V_0 + \lambda z\delta V_F\right].
\end{aligned}$$

Then (x, y, z) is a stationary symmetric equilibrium if

$$x \in BR(\Delta_L), \qquad y \in BR(\Delta_F), \qquad z \in BR(\Delta_0),$$

where

$$BR(\Delta) = \begin{cases} \{1\}, & \Delta > 0,\\ [0,1], & \Delta = 0,\\ \{0\}, & \Delta < 0. \end{cases}$$

This is a compact fixed point problem. A finite-horizon version is solved by backward induction. The infinite discounted version has stationary equilibria by the standard existence logic for discounted stochastic games after converting private signals into type-contingent actions.

6 Immediate comparative statics

Two inequalities are especially transparent.

For a leader with a release opportunity, acting is a defensive move. It matters only if the follower might also act. The leader's gain is

$$\Delta_L = \lambda y \left[\frac{s}{2} - k + \frac{\delta}{2}(V_L - V_F) \right].$$

So a leader defensively releases when the collision payoff plus half the discounted lead premium is nonnegative:

$$k \leq \frac{s}{2} + \frac{\delta}{2}(V_L - V_F).$$

If k is large, the incumbent prefers not to co-release even if it has a button.

For a follower, acting is offensive. If the leader is unlikely to act defensively, acting almost surely wins the lead. If the leader is likely to defend, the follower trades off the value of becoming leader against the collision haircut.

The no-leader state behaves like a symmetric preemption race. Acting can create the first leader, while waiting risks letting the opponent do so. The no-leader strategy z is therefore usually high unless collision costs are extreme and arrival hazards are high.

7 Numerical experiments

The following computations solve the fixed point approximately by grid search over mixed release probabilities. The goal is not numerical precision; it is to show the qualitative mechanics of the baseline model.

7.1 A representative parameter slice

Set $s = 1$ and $\delta = 0.95$. First take a moderate breakthrough arrival probability, $\lambda = 0.25$.

```
rows = [
    approximate_equilibrium(lam=0.25, k=k, grid_size=41)
    for k in [0.0, 0.25, 0.5, 1.0, 2.0, 5.0]
]
table = pd.DataFrame(rows)[
    ["k", "x_leader", "y_follower", "z_no_leader", "V_leader", "V_follower", "lead_premium"]
]
table.round(3)
```

	k	x_leader	y_follower	z_no_leader	V_leader	V_follower	lead_premium
0	0.00	1.0	1.0	1.0	10.604	9.396	1.208
1	0.25	1.0	1.0	1.0	10.292	9.083	1.208
2	0.50	1.0	1.0	1.0	9.979	8.771	1.208
3	1.00	0.0	1.0	1.0	10.476	9.524	0.952
4	2.00	0.0	1.0	1.0	10.476	9.524	0.952
5	5.00	0.0	1.0	1.0	10.476	9.524	0.952

At this arrival rate, the follower and no-leader release probabilities stay near one. The main margin is the incumbent's defensive release. When the co-location haircut is small, the leader is willing to press its own button to avoid being displaced. When the haircut is large, the leader waits and lets the challenger take the lead if the challenger has a release.

7.2 Arrival hazards and collision costs

Now vary the collision haircut k and compare a low-arrival environment to a high-arrival environment.

```
ks = np.linspace(0.0, 5.0, 41)
lambda_values = [0.10, 0.25, 0.50, 0.80]
records = []
for lam in lambda_values:
    for k in ks:
        records.append(approximate_equilibrium(lam=lam, k=float(k), grid_size=21))
df = pd.DataFrame(records)

fig, axes = plt.subplots(2, 2, figsize=(9.0, 6.0), sharex=True, sharey=True)
axes = axes.ravel()
for ax, lam in zip(axes, lambda_values):
    sub = df[df["lambda"] == lam]
    ax.plot(sub["k"], sub["x_leader"], label="leader $x$", linewidth=2)
    ax.plot(sub["k"], sub["y_follower"], label="follower $y$", linewidth=2)
    ax.plot(sub["k"], sub["z_no_leader"], label="no leader $z$", linewidth=2)
    ax.set_title(rf"$\lambda={lam}$")
    ax.set_ylim(-0.03, 1.03)
    ax.grid(alpha=0.25)
for ax in axes[2:]:
    ax.set_xlabel("collision haircut $k$")
for ax in axes[::2]:
    ax.set_ylabel("release probability")
```

```python
axes[0].legend(loc="lower left", fontsize=8)
plt.tight_layout()
plt.show()
```

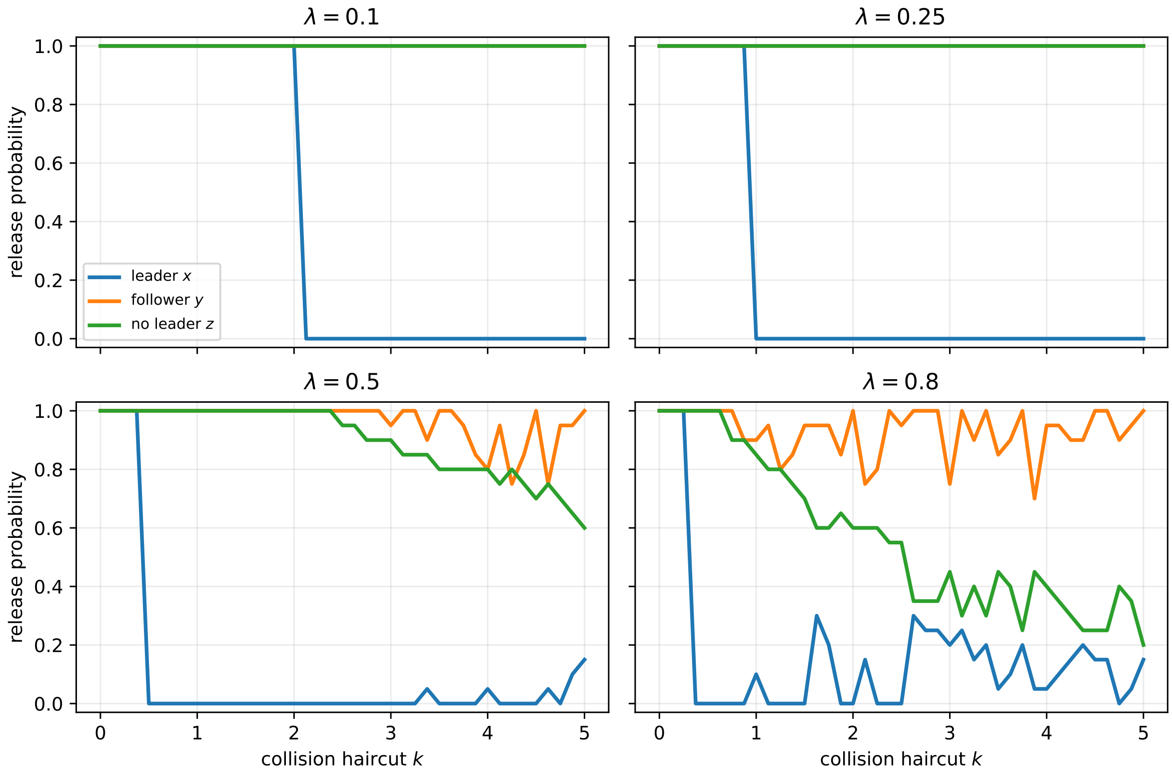

Figure 1: Approximate stationary release probabilities. x is leader defense, y is follower release, and z is no-leader release.

The pattern is intuitive:

- With rare breakthroughs, simultaneous release is unlikely, so even a sizable haircut does not immediately shut down defensive releases.
- With frequent breakthroughs, collision risk is high. The leader's defensive release probability falls quickly as k rises.
- The follower's release probability is high in most of the parameter space because a unilateral follower release captures the lead.
- In high-arrival/high-haircut regions, the no-leader state can involve mixed behavior: both firms want the first lead, but both dislike colliding.

The lead premium can also shrink or become negative. That does not mean the flow surplus is worthless. It means the leader is also the target. When arrival rates are high and collision costs are large, a follower has an offensive option while the leader faces displacement and collision risk.

```
fig, ax = plt.subplots(figsize=(7.5, 4.2))
for lam in lambda_values:
    sub = df[df["lambda"] == lam]
    ax.plot(sub["k"], sub["lead_premium"], label=rf"$\lambda={lam}$", linewidth=2)
ax.axhline(0, color="black", linewidth=1)
ax.set_xlabel("collision haircut $k$")
ax.set_ylabel(r"$V_L - V_F$")
ax.grid(alpha=0.25)
ax.legend()
plt.tight_layout()
plt.show()
```

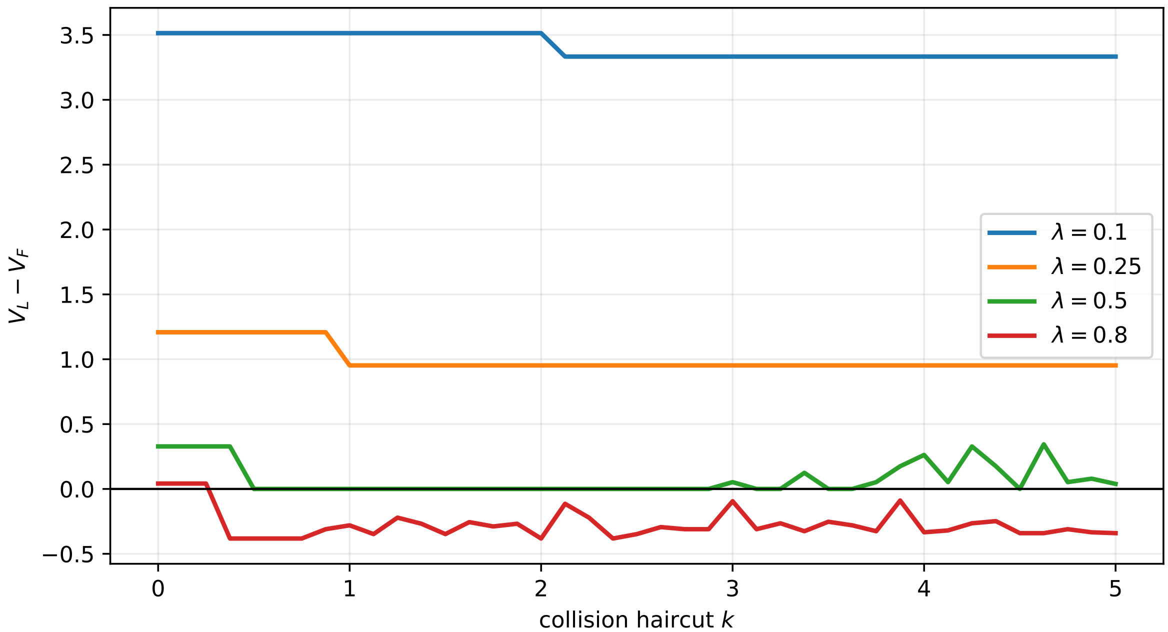

Figure 2: Approximate value of being leader rather than follower.

8 Persistent private buttons

The model above intentionally collapses the private threshold process into a fresh one-period opportunity. If the release button persists until pressed, the state must include hidden readiness.

Let

$$B_{it} \in \{0, 1\}$$

denote whether firm i currently has a private release button. A simple persistent law is

$$B_{i,t+1} = \begin{cases} 0, & a_{it} = 1, \\ 1, & a_{it} = 0,\ B_{it} = 1, \\ \text{Bernoulli}(\lambda_i), & a_{it} = 0,\ B_{it} = 0. \end{cases}$$

Equivalently, $B_{it} = 1$ once a latent capability process X_{it} crosses a threshold b_i, and action resets or transforms the process.

Now 'wait' is informative but not fully revealing. If a firm waits, observers update beliefs that the firm either had no button or had one but chose not to press it. A public belief

$$\mu_{it} = \Pr(B_{it} = 1 \mid \text{public history})$$

becomes part of the Markov state. A Markov perfect Bayesian strategy has the form

$$\sigma_i(a_i = 1 \mid \ell_t, B_{it}, \mu_{jt}),$$

and equilibrium consists of strategies, value functions, and Bayesian belief-updating rules. The Bellman equations are conceptually the same, but they live on a continuous belief state (μ_1, μ_2) unless one uses a finite grid or a special conjugate structure.

This persistent-button version is closer to the story "firms privately know when they can release." It is also exactly where the observed action data become hard to interpret: long waits may mean low idea arrival, high collision aversion, or strategic stockpiling.

9 What would be needed for empirical use?

For structural work, the minimal data-generating structure should specify:

- the calendar and surplus process s_t;
- the initial leader state;
- an arrival law for private release opportunities;
- whether opportunities persist;
- the tie continuation rule;
- whether release quality or release cost is observed;
- firm heterogeneity in λ_i, k_i, discounting, or surplus capture;
- what variation identifies arrival hazards separately from strategic waiting.

With only binary action histories, the model is weakly identified. A high observed waiting rate can come from low breakthrough hazards, large collision costs, patient firms, low surplus, or beliefs that the opponent is also ready. Useful extra variation would include release quality, internal capability milestones, exogenous shocks to deployment cost, market-wide surplus shocks, or public signals about competitor readiness.

10 Continuous-time analogue

The continuous-time version is a stochastic preemption game with private Poisson or threshold-crossing breakthroughs. A minimal analogue has private breakthrough intensities $\lambda_i(t)$ and release hazards chosen after a breakthrough. The current leader receives flow $s(t)$ until displaced. If both players place atoms at the same time, the collision payoff is $s(t)/2 - k$.

With independent Poisson arrivals and no strategic atoms, exact simultaneity has probability zero. Simultaneous release becomes important when players use atoms, deadlines, public news, synchronized evaluation windows, or common shocks. The continuous-time solution is usually written with Hamilton-Jacobi-Bellman or quasi-variational inequalities. The discrete-time model above is the same economic object without differential equations.

11 Bottom line

The clean relabelled workhorse is: **a discounted stochastic preemption game with private release opportunities and leader flow payoffs**.

The original verbal game is solvable only after adding the missing primitives. A particularly tractable baseline uses iid one-period private opportunities and stationary Markov perfect Bayesian equilibrium. In that model, followers usually release when able, incumbents defensively release only when collision costs are not too large, and high arrival rates make simultaneous-release risk central. The persistent-hidden-button version is the natural next step, but it is a belief-state game rather than a three-value fixed point.