Available online at www.sciencedirect.com

ScienceDirect

Journal of Economic Theory 212 (2023) 105702

JOURNAL OF Economic Theory

www.elsevier.com/locate/jet

The effects of rivalry on scientific progress under public vs private learning ☆

Heidrun Hoppe-Wewetzer [a,b], Georgios Katsenos [a], Emre Ozdenoren [c,b,*]

[a] *Leibniz University Hannover, Economics and Management, Germany*
[b] *CEPR, United Kingdom*
[c] *London Business School, United Kingdom*

Received 8 September 2021; final version received 5 July 2023; accepted 14 July 2023
Available online 20 July 2023

Abstract

We offer a model of scientific progress in which uncertainty resolves over time. We show that rivalry leads to less experimentation, extending results for preemption games to experimentation with uncertain outcomes. We compare experimentation duration and welfare when experimental outcomes are publicly versus privately observable. We show that public learning can generate more experimentation and higher welfare when uncertainty about the feasibility of a breakthrough is large; breakthroughs are rare even when they are feasible; and experiments produce results infrequently. Our results shed light on recent criticism of the science system.

JEL classification: D83; O31

Keywords: Stopping game; Experimentation; Learning; Preemption; Multiarmed bandit problem

☆ We are grateful to the editors and three anonymous referees for very valuable comments. We thank Catherine Bobtcheff, Marina Halac, Ehud Kalai, Peter A. Lawrence, Thomas Mariotti, Sven Rady, Yiman Sun, Stephan Thomsen, Julia Wirtz, and participants at the ESWC 2020 (Bocconi) for helpful comments on an earlier version of the paper entitled "Experimentation, Learning and Preemption". We thank especially Morton Kamien and Konstantin Wewetzer for discussions that inspired this work. Georgios Katsenos wishes to thank his early academic mentor Wilbur Jónsson. Financial support of the German Research Foundation (DFG HO 3814/1-1) is gratefully acknowledged.

* Corresponding author.

E-mail addresses: hoppe@mik.uni-hannover.de (H. Hoppe-Wewetzer), katsenos@mik.uni-hannover.de (G. Katsenos), eozdenoren@london.edu (E. Ozdenoren).

https://doi.org/10.1016/j.jet.2023.105702

1. Introduction

Scientific progress, according to Arrow (1969), is in the first instance the reduction in uncertainty: *"The product of a research and development effort is an observation on the world which reduces its possible range of variation"*. In fact, he argued that the information gain from an experiment might be more important than its concrete output. Challenging earlier models of research and development, Arrow calls for a more general formulation of knowledge production, encompassing situations where the probabilities of potential research outcomes are not known with certainty. Although uncertainty about outcomes is ubiquitous in science, very few formulations of this type have since been proposed in the literature (for an exception, see, e.g., Halac et al., 2017).

In this paper, we offer a model that combines uncertainty about the probabilities of research outcomes, as suggested by Arrow, with another typical feature of research activity: the competition to be first. Scientists seek to establish priority by being first to publish an advance in knowledge and are often concerned at being preempted in this by other scientists. Indeed, *"Since the earliest days of science, bragging rights to a discovery have gone to the person who first reports it"* (Fang and Casadevall, 2012).[1]

The main objective is to understand how the combination of learning about the distribution of research outcomes and preemption affects scientific progress and welfare. Our analysis also allows us to address a recent critique of the current science system: According to Lawrence (2016), the practice of university administrators to rank scientists against each other based on publications numbers, and allocate funds and jobs respectively, is impeding scientific progress by enhancing the importance attached to being first: *"All of us (...) focus our research to produce enough papers to compete and survive. Thus, projects are published as soon as possible and many therefore resemble lab reports rather than fully rounded and completed stories. (...) I think this emphasis on article numbers has helped make papers poorer in quality."* For biology and medicine, Broad (1981) observes that teams often settle for the *"least publishable unit"* - a practice that has come under fire for leading to research outcomes of lower quality overall. Adding to the criticism, the editors of *Nature* urged scientists conducting laboratory studies to take greater care in their work, citing several types of *"avoidable errors"*, in terms of both methodology and presentation, that diminish the quality of the published output (Nature Publishing Group, 2012). In response to the critique, Fang and Casadevall (2012) and, more recently, Stein and Hill (2021) have advocated a new science system that offers greater collegiality, freer sharing of information, and cooperation.

To set the stage, we study the extent of experimentation in a two-player stopping game and compare it to its counterpart in a setting without competition, which corresponds to the cooperative problem. As we show in our benchmark result, cooperation indeed always leads to more experimentation and value. However, competition is almost always an inherent feature of scientific inquiry. This raises the important question of whether transparency and sharing of research progress leads to more or less experimentation and value in a competitive setting with uncertain research outcomes. The answer is not immediate since there are competing forces. On the one hand, keeping research progress private might soften the competitive preemption threat (cf. Hopenhayn and Squintani, 2011), but on the other hand, public information may reduce the uncertainty about the likelihood of eventual success.

[1] For empirical evidence of the winner-takes-all rewards structure in science, see Hagstrom (1974), Newman (2009), and Sabatier and Chollet (2017).

To better understand this trade-off, we compare experimentation when the outcomes of the experiments are publicly versus privately observable. We show that, under certain conditions, public learning generates more experimentation and higher welfare. More precisely, we find that public learning tends to counteract the threat of preemption when uncertainty about the feasibility of a breakthrough is large, breakthroughs are rare even when they are feasible, and experiments frequently fail to produce results. In scientific research, we can approximate the probability of a breakthrough with the frequency of publishing a landmark paper, which appears to be quite low.[2] Hence, our result supports the views of Fang and Casadevall (2012) and Stein and Hill (2021), who are in favor of freer information sharing. Our findings may be surprising, particularly in the light of Hopenhayn and Squintani (2011), who show that secrecy may result in longer durations of experimentation by reducing the researcher's fear of being preempted. While there are several conflicting effects, we trace our results to the stronger ability to coordinate on the information obtained through experimentation when it is shared. This is one of the central insights of this paper.

Formally, we study a model in which two researchers running successive experiments decide at any point in time whether to stop and go forward with their best research finding thus far. Each experiment, with some probability, is successful, and the player receives a draw from some unknown distribution interpreted as the result of the experiment. With complementary probability, the experiment is unsuccessful and fails to produce any results. As we show later, the possibility of failed experiments distinguishes public and private learning.[3] The unknown distribution of draws remains fixed throughout the game, either producing low-value draws with certainty or randomizing between low- and high-value draws. We interpret a low-value draw as a mundane result and a high-value one as a breakthrough result from the project. To capture the uncertainty about the potential of the project, we assume the researchers do not know which is the true distribution, and they only share a prior belief about the feasibility of a high-value outcome. The competition is winner-takes-all, so, researchers have an incentive to stop preemptively and *"publish their partial findings quickly, rather than dropping the bombshell of a completely solved problem on their surprised colleagues"* (Hagstrom, 1974).[4]

We construct perfect Bayesian equilibria in symmetric threshold strategies. When the experimental outcomes are public, we establish the existence of equilibria in which the two players share common beliefs about the potential of the project and remain in the game until either a draw of high value occurs or their beliefs about the possibility of such a draw become too pessimistic. The latter event occurs when the total number of low-value draws exceeds a certain threshold, with the consequence that the players decide to stop simultaneously in equilibrium.

Our analysis in the case of private learning is complicated because of the complexity of the belief structure. Each player has to form beliefs regarding the draws his opponent has received. These beliefs and the player's own results determine in turn the player's belief about both the feasibility of a high-value outcome and the threat of preemption. In general, since the players' beliefs are private, it is difficult to track their evolution and, thus, to establish the existence of an equilibrium. The use of time as a public variable allows only for a partial simplification

[2] See, e.g., Bornmann et al. (2018).

[3] In fact, in natural science and many branches of social science, failure abounds (see, e.g., Mohs and Greig, 2017; Barwich, 2019 for empirical evidence). As Parkes (2019) notes: "*If we want to make new discoveries, that means taking a leap in the dark - a leap we might not take if we're too afraid to fail.*".

[4] We present the related literature in Section 7 where we classify existing models according to whether they deal with preemption, or uncertainty and learning about the distribution of research outcomes, or both.

of the belief structure because each player's beliefs about the number of low-value draws the other player has obtained depends on the number of low-value draws the player has himself obtained, as well as the other player's equilibrium strategy. Despite this complication, we are able to construct symmetric equilibria in strategies involving nonmonotone time-dependent thresholds: Each player experiments until he receives a high-value draw or accumulates too many low-value draws, although the threshold for the number of low value draws may vary non-monotonically over time.

Finally, we compare the length of experimentation and the players' total welfare under public versus private learning. Without the possibility of failed experiments, public and private learning are identical but otherwise they generate different outcomes. We find that public learning generates more experimentation than private learning, despite the higher threat of preemption, for a range of model parameters for arbitrary time horizons. Specifically, we find that if there is a lot of uncertainty about the feasibility of a breakthrough or if breakthroughs are rare even when feasible, then public learning generates more experimentation and higher welfare than private learning. These results provide testable implications of our model.

The paper is organized as follows. In Section 2, we present the model. In Section 3,we analyze the single-player case and cooperative benchmark. In Section 4, we analyze the two-player case under the assumption of public learning. In Section 5, we consider the case in which the two players cannot observe one another's draws. We provide a comparison between the two information settings in Section 6. In Section 7, we discuss how our results relate to the existing literature. We conclude in Section 8.

2. Model

Two players, 1 and 2, engage in a stopping game of successive experiments, taking place in discrete time periods $t = 1, \dots, T$. At the beginning of each period t, as long as the game continues, each player $i \in \{1, 2\}$ runs a new experiment. With probability $1 - r$, where $r \in (0, 1)$, player i's experiment is unsuccessful and fails to produce any valuable result. With probability r, the experiment is successful and provides new information about the common natural world. This is expressed by a draw $x_t^i \in \{L, H\}$ for player i in period t, where $0 < L < H$. That is, a successful experiment either provides some partial finding (of value L) or yields an important discovery (of value H). Incremental improvements over time are neglected in our formulation in order to sharpen the focus on the players' incentives to keep going, even though experiments may fail, in the hope of making a significant discovery.

An inherent feature of experimentation is the uncertainty regarding the potential outcomes of an experiment, which in our model is expressed by an uncertain distribution of the draws. Specifically, the values x_t^i are distributed according to either

$$x_t^i = \begin{cases} H, & \text{with probability } q; \\ L, & \text{with probability } 1 - q, \end{cases}$$

where $q \in (0, 1)$, or

$$x_t^i \equiv L.$$

The distribution is chosen randomly (by *nature*) at the beginning of the game, with probabilities p and $1 - p$ respectively, in a manner unobservable to the players, and remains the same throughout the game.[5] Conditional on the choice of distribution, the values x_t^i are independent across players and across periods. Thus, unless a draw of value H is obtained in an experiment, whether such an outcome is at all possible is unknown to the players.

We will consider two opposite cases regarding the observability of the players' experimentation outcomes: one in which each player can observe the draws of his opponent and the other in which each player can observe only his own draws.

At the end of each period t, each player i has to decide, after observing his own draw, x_t^i, and possibly his opponent's draw, x_t^j, whether to stop in that period or continue to period $t+1$. These actions are denoted by s or c, respectively. The two players make their decisions simultaneously and the game continues until at least one player decides to stop.

We assume that the experiments of the two players are directly competitive: the player who stops first receives a payoff equal to the value of his best past draw, while his opponent receives nothing. This winner-takes-all assumption seems particularly suited for a model of rivalry among scientists.[6] If both players decide to stop at the same time, with the same value, then we assume that only one of them – each with probability 1/2 – actually succeeds and becomes the first mover.[7] However, if the two players stop simultaneously with different values, then the player with the higher value receives his value in full whereas the other player gets zero. The two players discount time by a common rate $\delta \in (0, 1)$ and suffer no other cost for remaining active in the game.[8] Thus, to avoid trivial outcomes, we assume that each player can stop only after he has obtained at least one draw.

For each player i, a (private) history $h_t^i \in H_t^i$ at the time of his decision in period t consists of the following elements, depending on our observability assumption:

a. Player i's own past draws $x_\tau^i \in \{\emptyset, L, H\}$, for $\tau = 1, \ldots, t$, where $\emptyset$ denotes the occurrence of no draw;
b. Player j's past draws $x_\tau^j \in \{\emptyset, L, H\}$, for $\tau = 1, \ldots, t$, when draws are publicly observable;
c. Trivially, the two players' past decisions to continue, (c, c), for $\tau = 1, \ldots, t-1$.

A strategy of player i in period $t < T$ indicates whether the player stops or continues at the end of period t, for any possible time-t history. Hence, player i's strategy in period t is a function

$$\sigma_t^i : H_t^i \to \{s, c\},$$

under the restriction that $\sigma_t^i(h_t^i) = c$, if $h_t^i \in H_t^i$ is such that $x_\tau^i = \emptyset$ for all $\tau \leq t$; while player i's strategy for the entire game is a finite sequence of time-t strategies,

$$\sigma^i = \{\sigma_t^i\}_{t=1}^{T-1}.$$

[5] Note that players are sampling from the same distribution. This assumption is met, for example, when scientists seek to identify facts about the common natural world rather than to invent potentially different new technologies.

[6] See, for instance, Gaston (1973 [p.107]), Hagstrom (1974), Lawrence (2016) for empirical evidence. The assumption that preemption destroys all value to the second player simplifies the exposition, but is not crucial to our results. Our analysis would apply as long as the claim of L by one player destroys some nontrivial part of the value that the other player can claim.

[7] See Hoppe and Lehmann-Grube (2005) for a discussion of this tie-breaking rule in timing games.

[8] Our analysis extends with only slight modifications to the case in which there is a constant cost for each period a player is active. Since the presence of a discount factor suffices to make experimentation costly and to provide incentives to a player to stop experimenting even if he faces no preemption threat, we have chosen not to include such costs in our model.

We focus on pure strategies. Thus, each player i's strategy at time t partitions the set of the player's histories H_t^i into stopping and continuation regions, $\bar{H}_t^i$ and $H_t^i \setminus \bar{H}_t^i$.

Finally, our solution concept is that of perfect Bayesian equilibrium.

3. The cooperative benchmark

We start our analysis by examining the case in which experimentation is carried out by a single player, who performs one experiment in each period. We then modify this setting, by allowing two experiments to be performed in parallel within each period, so as to obtain the solution for the benchmark cooperative problem.

Clearly, the player will not stop before obtaining at least one draw and will not continue after obtaining a draw of H. Hence, the problem reduces to choosing whether to stop experimenting, claiming a value of L, or to continue at a cost of $(1-\delta)L$ for each additional period to potentially increase this value by $\delta(H-L)$.

The expected payoff from continuing to the next period depends on the player's belief about the distribution from which he draws. The player becomes more pessimistic that a draw of value H is feasible each time he receives a new draw of L. In particular, if the player has received $n \geq 1$ draws of L, then the player believes that he draws from the first distribution with probability

$$p\,(n) \;=\; \frac{(1-q)\; p\,(n-1)}{1-q\;\, p\,(n-1)}, \tag{1}$$

defined recursively, with $p\,(0) = p$. The sequence $\{p\,(n)\}_{n=0}^{\infty}$ is decreasing, since we have $p\,(n)\,/\,p\,(n-1) < 1$, for all $n \in \mathbb{N}$. Therefore, the expected value of staying in the game one more period weakly decreases as the game progresses.

Hence, for each period $t < T$, after having received $n_t \geq 1$ draws of L, the player will continue to period $t+1$ if and only if his expected one-step continuation payoff, discounted by δ, is larger than his stopping payoff, that is,

$$\delta\,[r\;p(n_t)\,q\;H \;+\; (1-r\;p(n_t)\,q)\,L] \;\geq\; L.$$

Thus, the optimal rule is to stop experimentation when $n_t \geq \hat{N}$ or $t = T$ and to continue otherwise, where

$$\hat{N} = \min\{n \in \mathbb{N} \;:\; \delta\, p(n)\, rq\,(H-L) \;<\; (1-\delta)\,L\}. \tag{2}$$

Finally, to obtain a proper cooperative benchmark for our analysis with two players, we modify the single-player case and allow the player to receive up to 2 draws in each period. This modification is necessary to account for the mere duplication of experiments with two players. In this case, given the player's beliefs $p(n_t)$ at the end of period t, the probability that the player obtains *at least* one draw of H in the period $t+1$ is

$$p^H(n_t) \;=\; p(n_t)\,[\,1-(1-rq)^2\,]. \tag{3}$$

Our previous analysis implies that two players who cooperate under an agreement to share information and, eventually, any value obtained will continue experimentation in periods $t = 1, \ldots, T-1$, until either at least one of them receives a draw of H or if they jointly obtain $n_t \geq N^*$ draws of L, where

$$N^* = \min\{n \in \mathbb{N} \;:\; \delta\, p^H(n)\,(H-L) \;<\; (1-\delta)\,L\}. \tag{4}$$

Since $p^H(n) \geq p(n)\,rq$, it follows that $N^* \geq \hat{N}$, reflecting the fact that two cooperating players experimenting in parallel are more likely to find H in the next period than one player experimenting alone, at the same cost (due to value depreciation).

In the sequel, we examine the impact of competition upon experimentation when outcomes are observed publicly or privately and compare these cases with each other as well as with the above cooperative benchmark.

4. Public learning

We now examine the two players' interaction. In this section, we assume that each player is fully informed of the experimental results of his rival. Players may have this information for various reasons. For example, they may be able to observe each other's experiments or there may be truthful communication between the players.

In this environment, in every period $t \geq 1$, the two players share *common beliefs* about the feasibility of an H outcome. If no draw of H has been obtained, these beliefs are expressed by the probability $p(n_t)$, where n_t is the total number of L draws obtained by the two players up to period t, determined recursively, according to equation (1) in the single-player problem. Hence, the probability that at least one draw of H is obtained by either player in the next period, if both players continue to it, is $p^H(n_t)$, defined by equation (3).

In the following, we construct a symmetric perfect Bayesian equilibrium in which experimentation terminates prior to the final period T if one or both players receive an H draw or if the total number of L draws reaches a certain threshold. In this equilibrium, like in the single player case, each player's expected gain from experimentation decreases as the number of L draws obtained (and jointly observed) by the two players increases.

First, suppose that by the time of the continuation or stopping decision in period t, each player has received at least one draw of L, that is, $n_t^i, n_t^j \geq 1$. In this case, the minimal number of L draws obtained by the two players such that a player will prefer to stop in period t rather than to continue to period $t+1$ and then surely stop if he knows that his opponent will also stop in period $t+1$ is

$$N_1 = \min\left\{ n \geq 2 : \, (1/2)\,\delta\,[\,p^H(n)(H-L) + L\,] < L \right\}. \tag{5}$$

Second, suppose that a single player has received all draws obtained by the end of period t. Then the minimal number of L draws such that this player will prefer to stop in period t rather than to continue to period $t+1$ and then surely stop if he knows that player j will stop as soon as he obtains a draw of L is[9]

$$N_2 = \min\left\{ n \geq 1 : \, (1/2)\,\delta\,[\,p^H(n)(H-L) + [1+(1-r)(1-p(n)rq)]\,L\,] < L \right\}. \tag{6}$$

By comparing the inequalities in the definition of N_1 and N_2, it can be shown that $N_1 \leq N_2$.

Furthermore, consider the *threshold strategy* $\sigma^* = \{\sigma_t^*\}_{t=1}^{T-1}$, prescribing to player i the following behavior in each period t:

- Player i stops in period t if he has obtained at least one draw and
  a. Player i has drawn H in some period $t' \leq t$; or

[9] The extra term in the left-hand-side of the inequality in the definition (6) of the threshold N_2 expresses the additional payoff that player i will receive in period $t+1$ in case player j does not obtain a draw in that period.

b. Player j has received a draw in some period $t' \leq t$, and $n_t^i + n_t^j \geq N_1$; or
c. Player j has received no draw in periods $t' \leq t$, and $n_t^i \geq N_2$.

- Otherwise, player i continues.

Clearly, the strategy σ^* is fully characterized by the thresholds N_1 and N_2, which remain constant over time.

Proposition 1. *The strategy profile* (σ^*, σ^*) *constitutes a perfect Bayesian equilibrium.*[10]

The equilibrium has a simple structure. The players remain in the game prior to the final period T until either a draw of high value occurs or their beliefs about the possibility of such a draw become too pessimistic. Since the players share common beliefs about the potential of the project, the latter event occurs when the *total* number of low-value draws exceeds a certain threshold. Consequently, in equilibrium, unless a draw of H is obtained, the players decide to stop simultaneously.

The game admits other equilibria in which the players stop experimenting after obtaining a total of $N' < N_1$ draws of L or after reaching a certain time T', where N' and T' are exogenously set. To see this, note that in such equilibria, because of the possibility of preemption, each player's decision to stop experimentation earlier forces his rival also to stop. However, experimentation resulting in more than N_1 or N_2 draws of L turns out to be impossible.

Proposition 2. *There exists no perfect Bayesian equilibrium involving experimentation that can generate more draws than the strategy* σ^*.

The following result compares public learning to the cooperative benchmark:

Proposition 3. *The maximal experimentation duration is longer in the case of two cooperating players than in any perfect Bayesian equilibrium under public learning.*

The proposition states that two players experimenting under an agreement to share information and value will search longer for H, in terms of the maximal number of experiments failing to find it, than two players sharing only information; that is, $N^* \geq N_2 \geq N_1$. Thus, the threat of preemption leads to a decrease in the total amount of experimentation, for a welfare loss.

5. Private learning

We now turn our attention to the case in which the two players cannot observe one another's experimental outcomes. Instead, in each period, each player has to form beliefs about the draws of his opponent, depending on the duration of experimentation, the stopping strategy his opponent has been using, and significantly, the draws he has received himself. Naturally, these beliefs affect the two players' continuation or stopping incentives, via their calculations about the likelihood of an H outcome as well as about the possibility that the other player stops in the current or next period.

In general, the beliefs of player i at time t take the form of a probability distribution over the feasible histories of the game, in particular, over the history components that are privately

[10] All proofs are in Appendix A.

observed by player j. In analyzing the stopping decision of player i in period t, when he has received no draw of H, we can assume that player j has received no draw of H either. Consequently, the beliefs of player i reduce to a probability distribution over the number of L draws, n_t^j, that player j has received up to period t.[11]

Since the probability of drawing L depends on the distribution from which the two players draw, player i's beliefs about n_t^j need to take into account his own private information, that is, the number n_t^i of L draws he has received.[12] In addition, player i needs to condition his beliefs upon any information he can infer from player j's decisions not to stop in any earlier period, given the strategy s^j.[13] The following result shows that the players' beliefs are positively correlated, that is, each player's beliefs about the draws of his opponent stochastically increase in the number of his own draws.

Lemma 1. *Suppose that player j follows the strategy s^j and that player i has obtained $n_t^i = n^i$ draws of L by period t. Then, at the end of period t, conditional on player j having received no draw of H, player i believes that $n_t^j = n^j$ with probability*

$$p_t(n^j, n^i, s^j) \;=\; \frac{h_t(n^j, s^j)\, r^{n^j}\, (1-r)^{t-n^j}\, [p(1-q)^{n^i+n^j} + (1-p)]}{\sum_{n=0}^{t} h_t(n, s^j)\, r^n\, (1-r)^{t-n}\, [p(1-q)^{n^i+n} + (1-p)]},$$

where $h_t(n^j, s^j) \leq \binom{t}{n^j}$ is the number of histories of player j consistent with $n_t^j = n^j$, the stopping constraints of strategy s^j, and the hypothesis that no draw of H has occurred.

In addition, for any $\tilde{n}^i > n^i$, the distribution $p_t(\cdot, \tilde{n}^i, s^j)$ first-order stochastically dominates the distribution $p_t(\cdot, n^i, s^j)$.

Given the symmetry of the game, we henceforth focus on equilibria in symmetric threshold strategies, that is, in which each player stops in period t if either he obtains a draw of H or the number of L draws he has received exceeds a certain threshold N_t, depending on that period. For such strategies, we can show that each player's beliefs are stochastically increasing in each threshold of his opponent:

Lemma 2. *Let s^j and $\hat{s}^j$ be two threshold strategies for player j such that $N_\tau^j \leq \hat{N}_\tau^j$ for all $\tau < t$. Then, for all n_t^i, the distribution $p_t(\cdot, n_t^i, \hat{s}^j)$ describing player i's beliefs about n_t^j at time t, conditional on player j having received no draw of H, first-order stochastically dominates the distribution $p_t(\cdot, n_t^i, s^j)$.*

Lemma 2 implies that, under private learning, in any period $t+1$, a player's belief that H is feasible, conditional on the game reaching that period, is monotonically decreasing in his rival's threshold in period t. Thus, under private learning, unlike the case of public learning, a player's

11 As Lemma 3 below will show, the timing of the players' draw arrivals is irrelevant in equilibrium.

12 For example, with a parameter $q \approx 1$, at the end of period $t = 1$, player i believes that H is feasible with probability approximately equal to p or 0, if, respectively, $n_t^i = 0$ or $n_t^i = 1$. Consequently, he believes that $n_t^j = 1$ with probability approximately equal to $(1-p)r$ or r, depending on whether $n_t^i = 0$ or $n_t^i = 1$.

13 In particular, if player j follows a strategy s^j characterized by stopping thresholds $\{N_t^j\}_{t=1}^{T-1}$, then player i will condition his beliefs at period t upon $n_{t'}^j < N_{t'}^j$ for all $t' < t$.

belief regarding the feasibility of H may update optimistically.[14] In addition, a change in the rival's threshold in period t affects a player's incentive to continue to period $t+1$ in both a positive and a negative manner, for an unclear overall effect. In fact, as Example 1 demonstrates below, a player may be more willing to continue if his rival adopts a lower threshold in the current period.

Our main result asserts the existence of a symmetric equilibrium in threshold strategies under a condition on the parameters of the model, ensuring that each player's expected payoff function from continuing or stopping at the end of each period $t < T$ satisfies a *single crossing property*.

Condition SC. The parameters δ, r, p, q, H, L and T are such that

$$p_t(N,1,\underline{s})\left[p(2N)\,[1-(1-rq)^2]-\frac{1-\delta}{\delta}\,\frac{L}{H-L}\right]\geq \sum_{n<N} p_t(n,1,\underline{s})\;p(n+1)\left[(1-rq)^2-(1-rq)^{2(T-t)}\right]$$

for all $N \leq t$ and $t < T$, where $\underline{s}$ is the strategy with thresholds $N_\tau = 1$ for $\tau < t-N$ and $N_\tau = \tau-(t-N)+1$ for $\tau \geq t-N$.

Using the expression for $p_t(n, n_t^i, s^j)$ in Lemma 1, with $n_t^i = 1$ and $s^j = \underline{s}$, the inequality in Condition SC becomes

$$r^N\,[p(1-q)^{N+1}+(1-p)]\left[p(2N)\,[1-(1-rq)^2]-\frac{1-\delta}{\delta}\,\frac{L}{H-L}\right]+ \sum_{n<N}\binom{N}{n} r^n\,(1-r)^{N-n}\,[p(1-q)^{n+1}+(1-p)]p(n+1) \times\left[(1-rq)^{2(T-t)}-(1-rq)^2\right]\geq 0$$

for all $N \leq t$ and $t < T$, which is easier to check.

The strategy $\underline{s}$ in Condition SC is "minimal" among the threshold strategies for which $n_t^j \geq N$ with positive probability; that is, if s^j is a threshold strategy such that $p_t(N,1,s^j) > 0$, then $N_\tau^j \geq N_\tau$ for all $\tau < t$. Therefore, by Lemma 2, the inequality in Condition SC extends to all such thresholds strategies s^j.[15]

Condition SC implies that player i's best-response in any period $t < T$ takes the form of a threshold N_t^i which is monotonically increasing in player j's threshold N_t^j.[16]

---

[14] For a simple example, suppose that at $t = 1$, player j always continues; and that at $t = 2$, player j continues only if he has received no draw. Then, if $n_1^i = n_2^i$, player i will be more optimistic about the feasibility of H in period $t = 2$ than in period $t = 1$.

[15] Notice that the right-hand-side in Condition SC is positive; so, in the left-hand-side, it follows that $p(2N)\,[1-(1-rq)^2]-\frac{1-\delta}{\delta}\,\frac{L}{H-L}\geq 0$. Thus, after a rearrangement of its terms, Condition SC requires that the expectation of an increasing function with respect to the distribution $p_t(\cdot,1,\underline{s})$ is positive, allowing Lemma 2 to apply.

[16] In particular, Condition SC is used in the proofs of Lemma 3 and Proposition 4 to show that each player's expected gain from continuing rather than stopping at the end of period t is decreasing in the number of L draws the player has observed. This property of monotone differences implies single crossing, which Milgrom and Shannon (1994) have shown to be necessary and sufficient for each player's best response in period t to be monotone (and thus, in our two-action setting, to be a threshold strategy).

For $t = T - 1$, the condition simplifies further to

$$\delta \left[p(2T) \left[1 - (1 - rq)^2\right] (H - L) + L \right] \geq L.$$

To describe how the condition is used in the argument, suppose that player j switches from a strategy $s_{T-1}(n^j_{T-1})$ of stopping in period $T-1$ to a strategy $\hat{s}_{T-1}(n^j_{T-1})$ of continuing in period $T-1$, for some n^j_{T-1}, with all other elements of his strategy remaining the same. Consequently, player i's payoff calculations involve a lower probability of player j stopping in period $T-1$ but also a lower expected payoff from experimentation, conditional on the game reaching period T, because of more pessimistic beliefs. Condition SC implies that player i's benefit from the switch in player j's strategy is greater when he continues to period T than when he stops in period $T-1$, for any number n^i_{T-1} of L draws that player i may have.

More generally, in any period $t < T$, suppose that player j has n^j_t draws of L and changes his strategy at time t from stopping to continuing and his continuation strategy from $\{s^j_\tau\}_{\tau=t+1}^{T-1}$ to $\{\hat{s}^j_\tau\}_{\tau=t+1}^{T-1}$.[17] Then player i's calculations about the benefits of further experimentation should involve not only more pessimistic beliefs, if the game reaches period $t+1$, but also a potential loss from the change in player j's continuation strategy. Condition SC requires that even under the worst-case scenario about the switch $\{s^j_\tau\}_{\tau>t}$ to $\{\hat{s}^j_\tau\}_{\tau>t}$, player i will benefit more from the change in player j's strategy in period t, if player i continues at time t rather than if he stops.

Although the condition is stronger than necessary, when it fails, a non-trivial symmetric equilibrium may not exist even for short time horizons. To see this, consider the following example:

Example 1. Let $\delta = 0.9$, $p = 0.8$, $q = 0.9$, $H = 8$, $L = 1$, and $T = 2$ (two periods). Then each player's strategy reduces to deciding whether to stop or to continue with one draw of L at the end of period $t = 1$. If $r \in (0.237, 0.242)$, then each player is better off stopping against an opponent who continues and continuing against an opponent who stops; therefore, there is no symmetric equilibrium.

In this example, q takes a relatively high value so that player j's decision to continue with one draw of L has a relatively large negative effect upon player i's beliefs about the feasibility of H, conditional on the game reaching period T.

The next lemma establishes the mutual optimality of the threshold strategies under Condition SC.

Lemma 3. *For any $T \in \mathbb{Z}^+$, if Condition SC holds, then each player i's best response to any threshold strategy $\{N^j_t\}_{t=1}^{T-1}$ of player j is also a threshold strategy $\{N^i_t\}_{t=1}^{T-1}$.*

The result of Lemma 3 is rather intuitive. With a higher number of L draws, player i becomes less willing to continue experimentation, for three reasons. First, independently of his opponent's presence, the extra draws of L have a negative effect upon player i's beliefs regarding the feasibility of H. Second, with another player experimenting in parallel, player i's pessimism about H is reinforced by the knowledge that the other player has not succeeded either. In particular, when player j will not stop unless he obtains H, player i's pessimism increases at a higher rate when

[17] In period $T-1$, a change in player j's continuation strategy is not possible.

he has received a higher number of L draws, independently of any preemption threat.[18] Third, considering also the opponent's stopping strategy, player i's fear of being preempted by the other player increases with each additional draw of L that he receives. In total, since the draws of L have only negative effects upon a player's expectations and payoffs, if player i is better off stopping with a certain number of L draws, then he will be better off stopping also with any higher number of such draws.

Suppose now that player j follows a strategy σ_j characterized by thresholds $\{N_t^j\}_{t=1}^{T-1}$. Then, at the end of each period t, player i's expected gain from continuing to period $t+1$ (and subsequently using his optimal continuation strategy) rather than stopping at period t, when he has obtained n_t^i draws of L, is

$$\Delta V_t \ = \ \Delta V_t(n_t^i \,|\, \sigma^j),$$

defined recursively by equations (A.1)–(A.6) in the proof of Lemma 3 (see Appendix A), with player i's beliefs about player j's draws being the ones induced from strategy σ^j via Lemma 1.

For any $T \in \mathbb{Z}^+$, a strategy σ with thresholds $\{N_t\}_{t=1}^{T-1}$ will be part of a symmetric equilibrium if, in each period $t < T$, we have

$$\Delta V_t(n_t^i \,|\, \sigma) \begin{cases} > 0 & \text{if } n_t^i < N_t, \\ \leq 0 & \text{if } n_t^i \geq N_t. \end{cases}$$

The following proposition asserts that such a symmetric equilibrium exists.

Proposition 4. *For any $T \in \mathbb{Z}^+$, if Condition SC holds, then there exists a symmetric perfect Bayesian equilibrium in threshold strategies $\{N_t\}_{t=1}^{T-1}$.*

The equilibrium strategies identified in Proposition 4 involve time-variant thresholds that may decrease and increase over time. While a player's best response within an examined period (i.e., the threshold to adopt in that period) is monotonic in the opponent's threshold, this does not imply that the player's entire continuation strategy (all thresholds in and after the period we consider) is monotonic in the opponent's threshold. To describe the way the thresholds N_t are determined, consider a player who has received $n_t^i = N$ draws of L by period t and who knows that his opponent will stop in that period if and only if he has also obtained $n_t^j \geq N_t^j = N$ draws of L. An increase in the number N has two effects upon the continuation incentives of that player: a positive one, stemming from the increase in N_t^j and the higher probability that his opponent will continue to the next period; and a negative one, stemming from the increase in n_t^i and the lower probability that H is feasible. As N increases, the second effect becomes more important. Eventually, either it comes to dominate the first effect, for a threshold $N_t \leq t+1$, or the two players choose always to continue experimenting for at least one more period.

6. Comparison of public and private learning

In this section, we compare the duration of experimentation and the players' total welfare under public and private learning. As we show, the players' total welfare is typically but not

[18] It is straightforward to calculate the probability that H is feasible, conditionally on n_t^i draws of L for player i and no draw of H for player j, and to show that the rate at which this probability decreases in the experimentation duration t is increasing in n_t^i.

always higher with longer experimentation. Our results indicate that public learning generates more experimentation when q is either low or high, r is low, and p is low. Private learning, on the other hand, generates more experimentation when q is intermediate, and r and p are high. In scientific research, there is often a great deal of uncertainty about the feasibility of a breakthrough (low p); breakthroughs are rare even when they are feasible (low q); and experiments frequently fail to produce results (low r). Hence, our findings suggest that public learning would generate more experimentation than private learning in scientific research. Throughout this section, when there are multiple equilibria, we focus on the equilibrium with the highest welfare.

As we noted before, the optimal experimentation duration and welfare are equal under both regimes when failed experiments are not possible. This is because when $r = 1$, under private learning, in each period, each player knows with certainty the number of L draws his opponent has received. However, when the arrival of draws is uncertain, i.e., for $r < 1$, public and private learning are no longer equivalent.

We first analyze the two-period case, in which equilibrium behavior can be completely characterized. In Proposition 7, we partially extend the two-period result to an arbitrary horizon by showing that public learning results in higher welfare relative to private learning under the parameter values for which the corresponding two-period result holds.

In the two-period case, Condition SC simplifies to the following inequality:

$$p(2)\,[1-(1-rq)^2]\,\frac{H-L}{L} \geq \frac{1-\delta}{\delta} \tag{7}$$

In addition, consider the following conditions upon a player's payoffs from continuing or stopping at the end of the first period, against an opponent who continues, depending on what the player can observe about his opponent's draw:

$$p(2)\,[1-(1-rq)^2]\,\frac{H-L}{L} \geq \frac{2-\delta}{\delta} \tag{8}$$

$$p(1)\,[1-(1-rq)^2]\,\frac{H-L}{L} + (1-r)[1-p(1)\,rq] < \frac{2-\delta}{\delta} \tag{9}$$

$$\begin{aligned}&[p_1(0,1)\,p(1)+(1-p_1(0,1))\,p(2)]\,[1-(1-rq)^2]\,\frac{H-L}{L}\\ &\qquad + p_1(0,1)\,(1-r)\,[1-p(1)\,rq] < \frac{2-\delta}{\delta}\end{aligned} \tag{10}$$

In these conditions, the probabilities $p(\cdot)$ and $p_1(0,1)$, expressing the players' beliefs, are defined, respectively, by equation (1) and Lemma 1.[19]

Proposition 5. *In the case of two periods, i.e., when $T = 2$, suppose condition (7) holds. Then the comparison of the most efficient equilibria under public and private learning depends on conditions (8)-(10):*

a. *If condition (8) holds, then public and private learning result in the same outcomes and payoffs, with each player continuing to period $T = 2$ unless he receives H.*

[19] Since $p(1) > p(2)$, the LHS in condition (8) is smaller than the LHS in condition (9), so that the two inequalities cannot hold simultaneously. In addition, the LHS in condition (10) is a convex combination of the LHS in conditions (8) and (9), weighted according to the players' beliefs $p_1(0,1)$ and $p_1(1,1)$; therefore, condition (10) must fail / hold respectively when condition (8) / (9) holds.

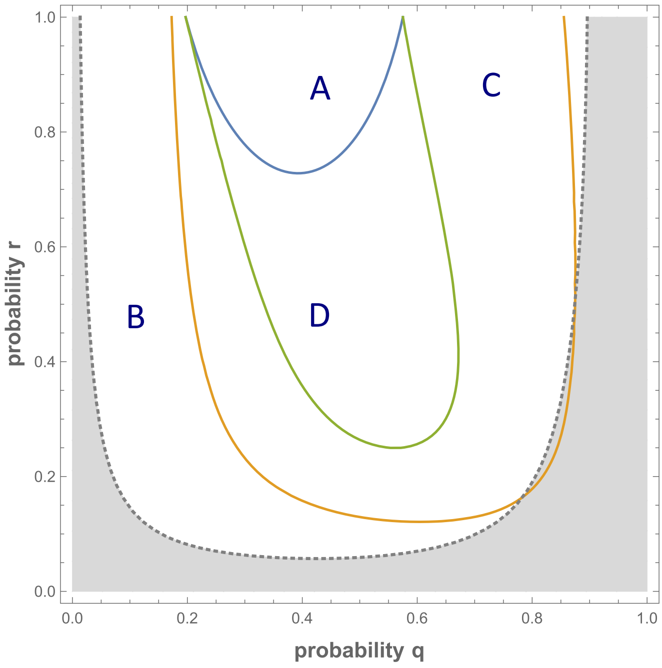

Fig. 1. For parameters $\delta = 0.9$, $H = 8$, $L = 1$, $p = 0.6$, condition (7) holds in the white area. Conditions (8) - (10) hold respectively inside the blue line (area A), outside the orange line (area B), outside the green line (areas B and C).

b. *If condition (9) holds, then public and private learning result in the same outcomes and payoffs, with each player stopping as soon as he receives a draw.*

c. *Otherwise, if conditions (8) and (9) do not hold, under public learning, the two players stop in period $t = 1$ if they both receive a draw of L; else, they continue to period $T = 2$. In this case, public learning generates more experimentation than private learning if and only if condition (10) holds. In addition, under condition (10), public learning results in higher expected payoffs.*

Fig. 1 illustrates the comparison of public and private learning when $T = 2$, for fixed parameters $\delta = 0.9$, $H = 8$, $L = 1$, $p = 0.6$ and variable probabilities r, q. To describe the graph, notice that the analysis of the players' incentives and strategies can be reduced to their decisions at $t = 1$ after obtaining a draw of L, depending under public learning on the information they observe about their opponent's draw. For each strategy profile, we indicate the combinations of r, q for which this profile constitutes an equilibrium. Under public learning there are three possible equilibria:

(i) Players always continue (*r, q in area A*).

(ii) Players stop if they both receive draws of L; otherwise, they continue (*r, q in areas C and D*).

(iii) Each player stops if he receives a draw of L regardless of the other player's draw (*r, q in area B*).

Under private learning, there are two possible equilibria:

(i) Players always continue (r, q *in areas A and D*).
(ii) Each player stops if he receives a draw of L (r, q *in areas B and C*).

Hence, in areas A or B, the length of experimentation is the same under public and private learning. In area D, private learning generates more experimentation than public learning. Finally, in area C, public learning generates more experimentation than private learning.[20]

To explain the intuition, note that private learning softens the threat of preemption, inducing players to stop later. Using Fig. 1, we see that private learning generates more experimentation than public learning when q has intermediate values and r is high enough. However, despite the possibility of preemption, we find that public learning generates more experimentation than private learning when q is either low or high enough. To see why this is the case, note that when q is low, i.e., breakthroughs are rare even when they are feasible, under private learning players stop with a single L because they believe that obtaining H with the next draw is very unlikely, so, the preemption motives dominate those of experimentation. On the other hand, when q is high, i.e., breakthroughs are rather frequent, obtaining an L leads players to update their beliefs drastically and believe that a breakthrough is not feasible (because if it were, they would have received an H with high probability given that q is high). This leads them to stop immediately. Under public learning, however, there is a range for the parameter q in which players would continue with a single L and stop only if they observe two Ls. Hence, when q is in this range, public learning generates more experimentation than private learning. Put differently, independent learning leads to coordination failures when players stop with a single L under private learning, but continue with a single L and stop if they both receive Ls under public learning. In addition, such coordination failures become more likely when r is low, i.e., experiments frequently fail to produce results. Indeed, for low values of r, public learning dominates private learning for all values of q.

Our analysis of the two-period case also reveals that the effect of softening preemption becomes stronger for a wider set of parameters when p gets higher, i.e., when it is more likely that a breakthrough is feasible. Graphically, in Fig. 1, as p increases, areas B+C contract, while areas A+D expand. We state this formally in the next proposition.

Proposition 6. *For $T = 2$, if public learning generates more experimentation and higher payoffs than private learning for some probability p, then it will generate weakly more experimentation and weakly higher payoffs for all probabilities $p' < p$.*

When public learning generates more experimentation, it necessarily results in higher welfare, as it is closer to the single-player optimum. It is interesting to notice, though, that public learning can result in higher welfare even in cases in which it generates less experimentation, if conditions (8)-(10) do not hold, where the solution to the cooperative problem is to experiment until obtaining $N^* = 2$ draws of L. For such parameters, the failure to aggregate the two players' information under private learning may result in excessive experimentation. The following example illustrates this possibility of excessive experimentation under private learning.

[20] As shown in the proof of Proposition 5, when condition (10) fails, in the areas A+D in Fig. 1, under private learning, an equilibrium exists even if condition (7) fails, with each player continuing to $T = 2$ unless he receives H in period $t = 1$. It is only for parameters for which (10) holds, in the areas B+C, that an equilibrium may not exist without condition (7) being satisfied.

Example 2. Let $\delta = 0.9$, $H = 8$, $L = 1$, $p = 0.9$, $q = 0.9$, $r = 0.1$, and $T = 2$. The cooperative solution is to keep experimenting until obtaining 2 draws of L. Under public learning, the two players stop at $t = 1$ if and only if they both obtain L draws; thus, the equilibrium achieves the optimal cooperative experimentation outcome. Under private learning, in equilibrium, players always continue to period $T = 2$ even if they each obtain an L draw. Hence, due to lack of coordination, it is possible that they continue experimenting beyond the cooperative stopping threshold. Thus, the expected duration/payoff of experimentation in the cooperative solution (and for public learning) is lower/higher than the expected duration/payoff under private learning.

For more than two periods, the comparison between public and private learning turns out to be complicated because of the large number of cases that need to be considered. However, as the following result shows, for parameters corresponding to area C in Fig. 1, public learning results in higher welfare relative to private learning, independently of the experimentation horizon T.

Proposition 7. *For $T \geq 2$, suppose that conditions (7) and (10) hold while condition (9) does not hold. Then public learning results in more experimentation and higher expected payoffs than private learning.*

In fact, for public learning to result in at least as efficient experimentation outcomes as private learning, it suffices that conditions (7) and (10) hold, that is, the parameters are in areas B and C in Fig. 1. Under these conditions, under private learning, there is a unique equilibrium, in which each player stops as soon as he receives one draw. The additional requirement that condition (9) does not hold restricts the set of parameters defined by conditions (7) and (10) to those in area C in Fig. 1, so that the players' thresholds under public learning are $N_2 \geq N_1 \geq 2$, for public learning to result in strictly better experimentation outcomes.

In Appendix B, we extend the analysis of the two-period problem under private learning by computing also non-efficient pure-strategy equilibria as well as mixed-strategy equilibria. In the latter, each player mixes between continuing and stopping if he has a draw of L at the end of the first period.

7. Related literature

Our paper is related to two bodies of work on experimentation, which are distinguished by the possibility of preemption and the presence of uncertainty regarding the distribution of potential outcomes. In preemption games, players decide when to terminate the game, given a first-mover advantage in the payoffs. They can seek to obtain a larger prize by moving late but also have the opportunity to accept a smaller prize, and by doing so, they prevent all others from obtaining any prize at all. In games of experimentation and learning, the players decide in each period whether to continue allocating resources to a risky project or to stop and exit for a safe option, according to what they can infer from the outcomes they have observed that far.

The first body of work features preemption in the sense that we just described, but does not deal with uncertainty and learning about the probability distribution of research outcomes. Hopenhayn and Squintani (2011) consider a preemption game in which two players randomly receive new information over time, interpreted as innovation increments. Players accumulate outcomes from a known distribution in their model. They find that private information about each player's state tends to soften the fear of being preempted, resulting in longer expected durations in equilibrium. This is in contrast to our findings. The reason is that, in our model, researchers not only learn about the threat of being preempted, but also about the probability distribution of research outcomes. We find that there are gains from making this information public that have no counterpart in their setup.

Bobtcheff et al. (2017) consider preemption in a model where two researchers privately have a breakthrough idea and decide how long to let the idea mature before disclosing it. However, the distribution of research outcomes is common knowledge. By contrast, our paper considers preemptive situations in which the feasibility of a high-value breakthrough is uncertain and focuses on learning about the distribution of outcomes and the effects of information exchange. Other preemption games in the context of research activity are investigated, for instance, by Lippman and Mamer (1993), Hoppe and Lehmann-Grube (2005). However, these studies consider preemption under deterministic payoffs.[21]

The other body of work deals with experimentation and learning in stopping games without the threat of preemption, as in the multi-armed bandit models (see, for instance, Keller et al., 2005). In these models, players must allocate resources to a risky project and a safe option. The risky project is characterized by uncertainty about the arrival rate of rewards, and players learn about this arrival rate over time by observing each other's actions and rewards. Private information in multi-armed bandit problems has been investigated by several authors, however, in these models there is no advantage from disclosing an experimentation result ahead of the opponent.[22]

Moscarini and Squintani (2010) consider a two-player experimentation model with learning about the arrival rate of an invention. In their setting, a player earns nothing when he stops before the invention arrives. Hence, preemption is not possible. By contrast, in our model, each player's beliefs regarding the position of his opponent are used to estimate not only the likelihood of achieving a high-value outcome but also the probability of being preempted with a low-value result.

Akcigit and Liu (2015) consider a model where two players begin experimenting with a risky arm that results in either a good outcome or a dead end. At any point, a player can privately and irreversibly switch to a safe arm. A good outcome from the risky arm is public, but a dead end is observed in private. Assuming that only a single player can obtain a reward from a given arm, the authors identify channels for inefficient experimentation. Aside from the different focus, the key difference between our paper and Akcigit and Liu (2015) stems from the lack of preemption in their framework. Without the threat of preemption, public experimentation is always superior to private experimentation.[23]

Heidhues et al. (2015) consider the possibility of communication via cheap talk in a multi-armed bandit model without preemption. Rosenberg et al. (2013), Dong (2021) and Wagner and Klein (2022) study the impact of private information about outcomes on welfare in two-armed bandit models without preemption. Margaria (2020) studies a two-player investment game with a second-mover advantage.

[21] Boyarchenko and Levendorskii (2014) examine preemption games with a single risky investment opportunity, but where learning about an uncertain distribution of outcomes is not an issue. Unlike us, they study the effects of players' asymmetry under jump-diffusion uncertainty.

[22] In our setting, the stopping and continuation decisions correspond, respectively, to settling for a sure arm and trying a stochastic arm. Note that in our model, a player's stopping decision affects the value of both arms for the other player.

[23] A second more technical difference concerns the evolution of beliefs in their setting where, for any strategies, players can only become pessimistic over time, unlike in our problem.

Building on the multi-armed bandit framework, Halac et al. (2017) study innovation contests when there is uncertainty about the feasibility of a successful innovation. There is a principal who designs a contest to maximize the probability of obtaining a successful innovation and several researchers who engage in costly experimentation for a fixed number of periods. The principal allocates a fixed prize among the researchers and chooses a prize-sharing scheme and a disclosure policy. Unlike in our model, preemption is not possible in their setting. This is because players cannot stop with anything less than a success, and even then, experimentation can continue after one of the players obtains success either because the contest is private or because there is equal sharing. By contrast, in our model, both private or public experimentation stop as soon as one of the players reveals either a low- or a high-value success which introduces fundamentally different learning and belief dynamics.

One paper that falls within the intersection of the two bodies of literature, dealing with preemption and learning about uncertain research outcomes, is Spatt and Sterbenz (1985). The authors show that preemption shortens experimentation. There are two crucial differences from our paper. First, in every period, there is a single public draw, and second, there are no failed experiments. Thus, there is no possibility of private learning in their setting, whereas our paper compares private and public learning. More recently, Bobtcheff et al. (2021) consider a preemption game with risky investment opportunity where players randomly receive a single perfectly informative private signal over time when the project is not profitable. In their model, learning is about bad news, and when learning is private, players delay investment to avoid the winner's curse. Thus a planner who would like to avoid unprofitable investment prefers private over public learning. By contrast, in our paper, learning is about the possibility of a breakthrough and we show that public learning can lead to longer experimentation and higher welfare. Hence, we view the papers as complementary.

8. Conclusion

We have examined the effects of rivalry upon experimentation and learning in a stopping game in which the players acquire information over time about the distribution of their potential payoffs. A key innovation in our setting is that experiments are not always successful and sometimes do not return any useful results.

Under the assumption of public observation of the players' experimentation results, we have constructed a perfect Bayesian equilibrium in threshold strategies. In this equilibrium, the two players continue experimenting, trying to obtain a high-value outcome, until their beliefs about its feasibility become too pessimistic. Because of the threat of preemption, the length of experimentation is shorter than socially optimal.

We have checked whether earlier results, showing that the threat of preemption is softened when information is kept private, carry over to preemption games with uncertainty and learning about research outcomes.[24] If players cannot observe one another's results, i.e., under private learning, they need to form beliefs about the experimentation outcomes of their rival and eventually about the feasibility of a high-value outcome. These beliefs turn out to be quite complex because they depend not only on the length of time the players have been experimenting but also on the number of successful experiments. Despite this complexity, we provide conditions for the existence of equilibria in strategies involving nonmonotone time-variant thresholds. Our anal-

[24] See Hopenhayn and Squintani (2011).

ysis reveals that private learning generates even shorter experimentation durations than public learning for a wide range of parameters.

We trace our findings to the players' inability to coordinate on their information under private learning: A player who does not observe his rival's experimentation results and, due to unsuccessful experimentation, does not himself have many results might still believe that his opponent has run many successful experiments and obtained more results. This situation would push the player to stop experimenting even earlier than under public observation of his rival's experimentation results. Overall, our paper sheds light on whether public or private experimentation generates longer experimentation horizons and greater value for scientists.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Appendix A. Proofs of results

Proof of Proposition 1. Consider the strategy $\sigma^* = \{\sigma_t^*\}_{t=1}^{T-1}$, described in Proposition 1. Arguing along the lines of the one-shot deviation principle, for (σ^*, σ^*) to constitute an equilibrium, we need to show that there is no continuation game such that player i can deviate profitably from the continuation strategy induced by σ^*. Suppose contrary and let t and h_t be the last period (within a finite time horizon) and one of the histories in that period in which player i can deviate profitably from the strategy σ^*. Thus, according to the contradiction hypothesis, σ^* is optimal for player i in all periods $t' > t$, against player j also following σ^*.

We split cases, depending on the history h_t.

Player i can clearly not gain from stopping after histories in which he has received no draw at all. In addition, player i cannot gain from continuing after histories in which he has already received a draw of H. Similarly, player i cannot deviate from σ^* profitably after histories in which his opponent has received a draw of H. So, it remains to check histories in which player i has received $n_t^i \geq 1$ draws of L and neither player has received a draw of H.

If $n_t^j \geq 1$ and $n_t = n_t^i + n_t^j < N_1$, then player i's payoff from following σ^* (thus, from continuing to period $t+1$ and then acting optimally, as implied by our contradiction hypothesis) weakly exceeds his payoff from continuing to period $t+1$ and then surely stopping, which is greater or equal than

$$(1/2)\,\delta\,[\,p^H(n_t)(H-L)+L\,] \;\geq\; L,$$

his payoff from stopping in period t.

On the other hand, if $n_t^j \geq 1$ and $n_t \geq N_1$, then player j will stop in period t, so, player i should also stop in that period.

If $n_t^j = 0$ and $n_t < N_2$, then player i's payoff from following σ^* (that is, from continuing to period $t+1$ and then acting optimally) weakly exceeds his payoff from continuing to period $t+1$ and then surely stopping, with player j also surely stopping if he receives a draw, which is

$$(1/2)\,\delta\,[\,p^H(n_t)(H-L)+[1+(1-r)(1-p(n_t)rq)]\,L\,]\;\geq\;L,$$

his payoff from stopping in period t.

If $n_t^j=0$ and $n_t\geq N_2$, if player i continues to period $t+1$, then player i will be best-off stopping in period $t+1$ (that is, following σ^* after period t); and player j will also stop in period $t+1$, if he receives a draw. Thus, player i's optimal deviation payoff is

$$(1/2)\,\delta\,[\,p^H(n_t)(H-L)+[1+(1-r)(1-p(n_t)rq)]\,L\,]\;<\;L,$$

his payoff from stopping in period t.

Having exhausted the cases, we have that player i has no profitable deviation from the strategy σ^*, contradicting our hypothesis. □

Proof of Proposition 2. We argue backwards, from period $t=T-1$ to period $t=1$, showing that at the end of each period t, the continuation game starting at that time cannot admit a pure-strategy equilibrium in which the players continue if the total number of L draws exceeds the thresholds N_1 and N_2 in the definition of σ^*.

In period $t=T-1$, when $n_t^j\geq 1$ and $n_t^i+n_t^j\geq N_1$, even if player j is willing to continue to period T, the inequality in the definition (5) of the threshold N_1 implies that player i is better-off stopping in period $T-1$. Thus, there is no equilibrium in which a player might continue to period T when $n_t^i+n_t^j\geq N_1$. In addition, when $n_t^j=0$ and $n_t^i\geq N_2$, the inequality in the definition (6) of the threshold N_2 implies that player j is better-off stopping in period $T-1$.

In period $t=T-2$, when $n_t^j\geq 1$ and $n_t^i+n_t^j\geq N_1$, even if player j is willing to continue to period T, player i knows that the game will surely end in the next period. Therefore, player i's continuation and stopping payoff calculations in period $T-2$ are identical to those in period $T-1$, implying again that there is no equilibrium in which the players might continue to period $T-1$ when $n_t^i+n_t^j\geq N_1$. For the same reason, the knowledge that the game will not continue beyond $T-1$, player i is better-off stopping in period $T-2$, when $n_t^j=0$ and $n_t^i\geq N_2$.

Reiterating the last argument for $t=T-3,\ldots,1$, noticing that in each period t, when $n_t^j\geq 1$ and $n_t^i+n_t^j\geq N_1$ or when $n_t^j=0$ and $n_t^i\geq N_2$, player i is forced to treat the continuation game as a two-period game, it follows that there is no equilibrium in which the game continues after histories in which the strategy σ^* dictates stopping. □

Proof of Proposition 3. Comparing the inequalities in (4) and (6), defining the thresholds N^* and $N_2\geq N_1$, we find that a player's gain from continuing experimenting for exactly one more period is larger when he is alone, so that $N^*\geq N_2\geq N_1$. □

Proof of Lemma 1. At the end of period t, consider the joint event in which the two players have observed respectively histories h_t^i and h_t^j involving n_t^i and n_t^j draws of L and no draw of H. The probability of this event is

$$P(h_t^i,h_t^j)\;=\;r^{n_t^i+n_t^j}\,(1-r)^{2t-n_t^i-n_t^j}\,[\,p(1-q)^{n_t^i+n_t^j}+(1-p)\,]$$

Aggregating over all time-t histories h_t^j involving n_t^j draws of L, no draw of H, and satisfying the continuation constraints of the strategy s^j for all periods up to time $t-1$, we get

$$P(h_t^i,n_t^j,s^j)\;=\;h_t(n_t^j,s_t^j)\,r^{n_t^i+n_t^j}\,(1-r)^{2t-n_t^i-n_t^j}\,[\,p(1-q)^{n_t^i+n_t^j}+(1-p)\,],$$

where $h_t(n_t^j, s_t^j) \leq \binom{t}{n_t^j}$ is the total number of such histories.

Therefore, player i's belief that $n_t^j = n^j$ is given by the conditional probability

$$p_t(n_t^j, n_t^i, s^j) = P(n_t^j \mid h_t^i, s^j) = \frac{P(h_t^i, n_t^j, s^j)}{\sum_{n=0}^t P(h_t^i, n, s^j)}$$
$$= \frac{h_t(n_t^j, s^j)\, r^{n_t^j}\, (1-r)^{t-n_t^j}\, [p(1-q)^{n_t^i+n_t^j} + (1-p)]}{\sum_{n=0}^t h_t(n, s^j)\, r^n\, (1-r)^{t-n}\, [p(1-q)^{n_t^i+n} + (1-p)]},$$

with the second equality being obtained by canceling equal terms.

To explore the monotonicity of the beliefs $p_t(n_t^j, n_t^i, s^j)$ with respect to the variable n_t^i, notice that

$$\frac{dp_t}{dn_t^i,}(n_t^j, n_t^i, s^j) =$$
$$\frac{\ln(1-q)\; h_t(n_t^j, s^j)\, r^{n_t^j}\, (1-r)^{t-n_t^j}}{\left(\sum_{n=0}^t h_t(n, s^j)\, r^n\, (1-r)^{t-n}\, [p(1-q)^{n_t^i+n} + (1-p)]\right)^2}$$
$$\times \sum_{n=0}^{t} h_t(n, s^j)\; r^n\, (1-r)^{t-n}\; p\,(1-p)\,(1-q)^{n_t^i}\, [(1-q)^{n_t^j} - (1-q)^n]$$

Therefore, since $\ln(1-q) \leq 0$,

$$\frac{dp_t}{dn_t^i,}(n_t^j, n_t^i, s^j) \gtreqless 0 \iff \sum_{n=0}^{t} h_t(n, s^j)\; r^n\, (1-r)^{t-n}\, [(1-q)^{n_t^j} - (1-q)^n] \lesseqgtr 0,$$

The sum is independent of n_t^i, decreasing in n_t^j, positive for $n_t^j = 0$, negative for $n_t^j = t$. Hence, for every t and s^j, there is a value $\bar{n}_t^j$ such that

$$\frac{dp_t}{dn_t^i,}(n_t^j, n_t^i, s^j) \gtreqless 0 \iff n_t^j \gtreqless \bar{n}_t^j$$

Let $\tilde{n}_t^i > n_t^i$. To show that

$$\sum_{n_t^j=0}^{n} [\, p_t(n_t^j, n_t^i, s^j) - p_t(n_t^j, \tilde{n}_t^i, s^j)\,] \geq 0, \quad \text{for all } n = 0, 1, \ldots t,$$

as required for first-order stochastic dominance, notice that

$$p_t(n_t^j, \tilde{n}_t^i, s^j) \gtreqless p_t(n_t^j, n_t^i, s^j) \iff n_t^j \gtreqless \bar{n}_t^j.$$

Therefore, the sum is positive for values $n \leq \bar{n}_t^j$. For values $n \geq \bar{n}_t^j$, we have

$$\sum_{n_t^j=0}^{n} [\, p_t(n_t^j, n_t^i, s^j) - p_t(n_t^j, \tilde{n}_t^i, s^j)\,] = - \sum_{n_t^j=n+1}^{t} [\, p_t(n_t^j, n_t^i, s^j) - p_t(n_t^j, \tilde{n}_t^i, s^j)\,]$$

so that again the sum is positive, as required. □

Proof of Lemma 2. Since first-order stochastic dominance is a transitive relation, so that our argument can proceed from s^j to $\hat{s}^j$ in a threshold-by-threshold manner, it suffices to show the result for strategies s^j and $\hat{s}^j$ such that $N_\tau^j = \hat{N}_\tau^j$, for $\tau \neq t_0$, and $N_\tau^j < \hat{N}_\tau^j$, for $\tau = t_0$, for some time $t_0 < t$.

Given two threshold strategies s^j and $\hat{s}^j$ that differ only at time $t_0 < t$, with $N_{t_0}^j < \hat{N}_{t_0}^j$, by Lemma 1, for all $M \le t$, we have

$$P[n_t^j \le M \,|\, n_t^i, \hat{s}^j] - P[n_t^j \le M \,|\, n_t^i, s^j] = \sum_{m=0}^{M} \left[\frac{h_t(m, \hat{s}^j)\, \bar{p}(m, n_t^i)}{\sum_{n=0}^{t} h_t(n, \hat{s}^j)\, \bar{p}(n, n_t^i)} - \frac{h_t(m, s^j)\, \bar{p}(m, n_t^i)}{\sum_{n=0}^{t} h_t(n, s^j)\, \bar{p}(n, n_t^i)} \right],$$

with the expression $\bar{p}(m, n_t^i) = r^m (1-r)^{t-m} [p(1-q)^{n_t^i+m} + (1-p)]$ being used to simplify the notation. Therefore, for all $M \le t$,

$$P[n_t^j \le M \,|\, n_t^i, \hat{s}^j] - P[n_t^j \le M \,|\, n_t^i, s^j] \le 0$$

as required for the result, if and only if

$$\sum_{m=0}^{M} \sum_{n=0}^{t} \bar{p}(m, n_t^i)\, \bar{p}(n, n_t^i) \left[h_t(m, \hat{s}^j)\, h_t(n, s^j) - h_t(m, s^j)\, h_t(n, \hat{s}^j) \right] \le 0$$

or, after canceling equal terms, if and only if

$$\sum_{m=0}^{M} \sum_{n=M+1}^{t} \bar{p}(m, n_t^i)\, \bar{p}(n, n_t^i) \left[h_t(m, \hat{s}^j)\, h_t(n, s^j) - h_t(m, s^j)\, h_t(n, \hat{s}^j) \right] \le 0$$

Therefore, it suffices to show that

$$h_t(m, \hat{s}^j)\, h_t(n, s^j) - h_t(m, s^j)\, h_t(n, \hat{s}^j) \le 0,$$

for all $m, n \le t$ such that $m \le M < n$.

Notice that for all strategies s with thresholds $\{N_\tau\}_{\tau=1}^{t-1}$ and any time $t_0 < t$, we have

$$h_t(k, s) = \sum_{l=0}^{k} h'_{t_0}[l, (N_\tau)_{\tau=1}^{t_0}]\, h_{t-1-t_0}[k-l, (N_\tau - l)_{\tau=t_0+1}^{t-1}]$$

where $h'_{t_0}[l, (N_\tau)_{\tau=1}^{t_0}]$ is the number of player j's histories at the end of period t_0 such that player j has received l draws of L and no draw of H and such that $n_\tau^j < N_\tau$ for all $\tau \le t_0$.

Therefore, it suffices to show that

$$\begin{aligned} &\sum_{k=0}^{m} h'_{t_0}[k, (\hat{N}_\tau^j)_{\tau=1}^{t_0}]\, h_{t-1-t_0}[m-k, (\hat{N}_\tau^j - k)_{\tau=t_0+1}^{t-1}] \\ &\qquad \times \sum_{l=0}^{n} h'_{t_0}[l, (N_\tau^j)_{\tau=1}^{t_0}]\, h_{t-1-t_0}[n-l, (N_\tau^j - l)_{\tau=t_0+1}^{t-1}] - \\ &\qquad \sum_{k=0}^{m} h'_{t_0}[k, (N_\tau^j)_{\tau=1}^{t_0}] h_{t-1-t_0}[m-k, (N_\tau^j - k)_{\tau=t_0+1}^{t-1}] \\ &\qquad \times \sum_{l=0}^{n} h'_{t_0}[l, (\hat{N}_\tau^j)_{\tau=1}^{t_0}]\, h_{t-1-t_0}[n-l, (\hat{N}_\tau^j - l)_{\tau=t_0+1}^{t-1}] \le 0 \end{aligned}$$

Since $\hat{N}_\tau^j = N_\tau^j$, for all $\tau > t_0$, this reduces to showing (after again canceling equal terms) that

$$\sum_{k=0}^{m} \sum_{l=m+1}^{n} h_{t-1-t_0}[m-k, (N_\tau^j - k)_{\tau=t_0+1}^{t-1}]\, h_{t-1-t_0}[m-l, (N_\tau^j - l)_{\tau=t_0+1}^{t-1}] \times \begin{bmatrix} h'_{t_0}[k, (\hat{N}_\tau^j)_{\tau=1}^{t_0}]\, h'_{t_0}[l, (N_\tau^j)_{\tau=1}^{t_0}] - \\ h'_{t_0}[k, (N_\tau^j)_{\tau=1}^{t_0}]\, h'_{t_0}[l, (\hat{N}_\tau^j)_{\tau=1}^{t_0}] \end{bmatrix} \leq 0$$

for all $m, n \leq t$ such that $m \leq M < n$.

For $m < N_{t_0}^j$, we have $h'_{t_0}[k, (\hat{N}_\tau^j)_{\tau=1}^{t_0}] = h'_{t_0}[k, (N_\tau^j)_{\tau=1}^{t_0}]$, for all $k \leq m$, so that the inequality follows from the fact that $h'_{t_0}[l, (N_\tau^j)_{\tau=1}^{t_0}] \leq h'_{t_0}[l, (\hat{N}_\tau^j)_{\tau=1}^{t_0}]$, for all $l \geq 0$.

Finally, for $m \geq N_{t_0}^j$, we have $h'_{t_0}[l, (N_\tau^j)_{\tau=1}^{t_0}] = 0$, for all $l \geq m+1$, so that the expression on the left-hand-side of the inequality involves only non-positive terms. □

Proof of Lemma 3. We argue by means of backwards induction, in periods $T-1, T-2, \ldots, 1$, showing in each period, first, that player i's optimal strategy at the end of the period takes the form of a threshold rule; and second, that player i's expected payoff from following his optimal strategy is decreasing in the number of L draws he has obtained that far, a result to be used in the next step of the induction.

Throughout our argument we condition on player j having obtained no draw of H by the time of player i's decision; otherwise, player i's decision is irrelevant for his payoff. For the sake of brevity, we drop this condition from our notation.

Given any $T \in \mathbb{Z}^+$, suppose that player j's strategy s^j is such that he stops in periods $t < T$ if and only if $n_t^j \geq N_t^j$, for some sequence of thresholds $\{N_t^j\}_{t=1}^{T-1}$.

Moving backwards in the periods of the game, suppose that player i has obtained $n_{T-1}^i > 0$ draws of L by the end of period $T-1$.[25] Then player i's expected payoff from continuing to the last period T, conditionally on player j having obtained n_{T-1}^j draws of L and on the game actually reaching period T, is

$$U_T(n_{T-1}^i | n_{T-1}^j, s^j) = \begin{cases} \frac{1}{2}\delta\,[p^H(n_{T-1}^i + n_{T-1}^j)(H-L) + L], & n_{T-1}^j > 0; \\ \frac{1}{2}\delta\,[p^H(n_{T-1}^i)(H-L) + L] + \\ \frac{1}{2}\delta\,[1 - r\,p(n_{T-1}^i)\,q]\,(1-r)\,L, & n_{T-1}^j = 0. \end{cases} \tag{A.1}$$

Therefore, conditionally on n_{T-1}^j, player i's expected gain from continuing to period T instead of stopping in period $T-1$ is

$$\Delta V_{T-1}(n_{T-1}^i \,|\, n_{T-1}^j, s^j) = \begin{cases} -L/2, & n_{T-1}^j \geq N_{T-1}^j; \\ U_T(n_{T-1}^i \,|\, n_{T-1}^j, s^j) - L, & n_{T-1}^j < N_{T-1}^j. \end{cases}$$

Finally, player i's (unconditional) expected gain from continuing instead of stopping is

$$\Delta V_{T-1}(n_{T-1}^i \,|\, s^j) = \sum_{n_{T-1}^j=0}^{T-1} p_{T-1}(n_{T-1}^j, n_{T-1}^i, s^j)\, \Delta V_{T-1}(n_{T-1}^i \,|\, n_{T-1}^j, s^j). \tag{A.2}$$

[25] When $n_{T-1}^i = 0$, player i must continue into period T independently of h_{T-1}^j.

Under Condition SC, the function $\Delta V_{T-1}(\cdot\,|\,\cdot, s^j)$ is decreasing in n^j_{T-1}.[26] In addition,

$$\Delta V_{T-1}(n^i_{T-1}\,|\,0, s^j) = (1/2)\delta\, p(n^i_{T-1})\, rq\,[(2-rq)H - (3-r-rq)L] - L + (1/2)\delta\,(2-r)L.$$

Therefore, for parameters $H/L < (3-r-rq)/(2-rq)$, we have $\Delta V_{T-1}(n^i_{T-1}\,|\,0, s^j) < 0$, so that $\Delta V_{T-1}(n^i_{T-1}\,|\,n^j_{T-1}, s^j) < 0$, for all $n^i_{T-1} \geq 1$, $n^j_{T-1} \geq 0$. In this case, player i's expected gain from continuing is $\Delta V_{T-1}(n^i_{T-1}\,|\,s^j) < 0$, for all $n^i_{T-1} \geq 1$, implying that player i is best-off stopping if he has at least one draw of L. Otherwise, for parameters $H/L \geq (3-r-rq)/(2-rq)$, the function $\Delta V_{T-1}(\cdot\,|\,\cdot, s^j)$ is decreasing also in n^i_{T-1}. In this case, for $\tilde{n}^i_{T-1} > n^i_{T-1}$, we have

$$\begin{aligned}\Delta V_{T-1}(\tilde{n}^i_{T-1}\,|\,s^j) &= \sum_{n^j_{T-1}=0}^{T-1} p_{T-1}(n^j_{T-1}, \tilde{n}^i_{T-1}, s^j)\,\Delta V_{T-1}(\tilde{n}^i_{T-1}\,|\,n^j_{T-1}, s^j)\\ &\leq \sum_{n^j_{T-1}=0}^{T-1} p_{T-1}(n^j_{T-1}, n^i_{T-1}, s^j)\,\Delta V_{T-1}(n^i_{T-1}\,|\,n^j_{T-1}, s^j)\\ &= \Delta V_{T-1}(n^i_{T-1}\,|\,s^j),\end{aligned}$$

with the inequality being obtained from the fact that the probability distribution $p_{T-1}(\cdot, \tilde{n}^i_{T-1}, s^j)$ first-order stochastically dominates the distribution $p_{T-1}(\cdot, n^i_{T-1}, s^j)$. Hence, player i's incentive to continue to period T is decreasing in the number n^i_{T-1} of L draws he has received, implying that his best response in period $T-1$ takes the form of a threshold rule, N^i_{T-1}.

To complete the first step of the induction, notice that player i's expected payoff from choosing to continue to period T,

$$V^c_{T-1}(n^i_{T-1}\,|\,s^j) \;=\; \sum_{n^j_{T-1}=0}^{N^j_{T-1}-1} p_{T-1}(n^j_{T-1}, n^i_{T-1}, s^j)\, U_T(n^i_{T-1}\,|\,n^j_{T-1}, s^j)$$

is decreasing in n^i_{T-1}, since the distribution $p_{T-1}(\cdot, n^i_{T-1}, s^j)$ is first-order stochastically increasing in n^i_{T-1} and the payoff $U_T(n^i_{T-1}\,|\,n^j_{T-1}, s^j)$ is decreasing in n^i_{T-1} and n^j_{T-1}. In addition, player i's payoff from stopping in period $T-1$,

$$V^s_{T-1}(n^i_{T-1}\,|\,s^j) \;=\; (L/2) \;+ \sum_{n^j_{T-1}=0}^{N^j_{T-1}-1} p_{T-1}(n^j_{T-1}, n^i_{T-1}, s^j)\;(L/2),$$

is also decreasing in n^i_{T-1}, because of stochastic dominance. Therefore, player i's optimal payoff at the end of period $T-1$,

$$V^*_{T-1}(n^i_{T-1}\,|\,s^j) \;=\; \max\{\, V^c_{T-1}(n^i_{T-1}\,|\,s^j),\; V^s_{T-1}(n^i_{T-1}\,|\,s^j)\,\} \tag{A.3}$$

[26] For all n^i_{T-1}, since the probability $p^H(n^i_{T-1} + n^j_{T-1})$ is decreasing in n^j_{T-1}, the payoff $U_T(n^i_{T-1}\,|\,n^j_{T-1}, s^j)$ is also decreasing in n^j_{T-1}. Condition SC ensures that $U_T(n^i_{T-1}\,|\,n^j_{T-1}, s^j) - L > -L/2$, for all $n^j_{T-1} < N^j_{T-1}$, for all N^j_{T-1}.

is decreasing in n^i_{T-1}.

Proceeding to periods $t = T-2, T-3, \dots, 1$, suppose that player i's optimal continuation strategy in period $t+1$ takes the form of a threshold rule $\{N^i_\tau\}_{\tau=t+1}^{T-1}$, depending only on the strategy s^j; and that his optimal payoff at the end of period $t+1$,

$$V^*_{t+1}(n^i_{t+1} \,|\, s^j) \;=\; V_{t+1}[n^i_{t+1} \,|\, s^j, (N^i_\tau)_{\tau=t+1}^{T-1}]$$

is decreasing in n^i_{t+1} (induction hypothesis).

At the beginning of period $t+1$, player i's expected payoff from drawing in that period and then following the optimal continuation strategy $\{N^i_\tau\}_{\tau=t+1}^{T-1}$ is

$$\begin{aligned} &U^*_{t+1}(n^i_t \,|\, s^j) = U_{t+1}[n^i_t \,|\, s^j, (N^i_\tau)_{\tau=t+1}^{T-1}] \\ &= \hat{p}_H(n^i_t \,|\, s^j)\,(1/2)\,H + \;[1-\hat{p}^H_t(n^i_t \,|\, s^j)]\;\hat{p}^L_t(n^i_t \,|\, s^j)\;V^*_{t+1}(n^i_t+1 \,|\, s^j) \\ &\quad + [1-\hat{p}^H_t(n^i_t \,|\, s^j)]\,[1-\hat{p}^L_t(n^i_t \,|\, s^j)]\;V^*_{t+1}(n^i_t \,|\, s^j) \end{aligned} \tag{A.4}$$

where

$$\hat{p}^H_t(n^i_t \,|\, s^j) \;=\; \sum_{n^j_t=0}^{t} p'_t(n^j_t, n^i_t, s^j)\, p_H(n^j_t + n^i_t)$$

is player i's belief at the beginning of period $t+1$ that at least one draw of H will be obtained in that period,

$$p^L_t(n^i_t \,|\, s^j) \;=\; \sum_{n^j_t=0}^{t} p'_t(n^j_t, n^i_t, s^j)\, \frac{[\,1-p(n^j_t+n^i_t)+p(n^j_t+n^i_t)\,(1-q)\,(1-rq)\,]\,r}{1-p(n^j_t+n^i_t)+p(n^j_t+n^i_t)\,(1-rq)^2}$$

is player i's belief at the beginning of period $t+1$ that he will draw L in that period, conditional on neither player drawing H, with

$$p'_t(n^j_t, n^i_t, s^j) \;=\; \frac{h'_t(n^j_t, s^j)\, r^{n^j_t}\,(1-r)^{t-n^j_t}\,[p(1-q)^{n^i_t+n^j_t}+(1-p)]}{\sum_{n=0}^{t} h'_t(n, s^j)\, r^n\,(1-r)^{t-n}\,[p(1-q)^{n^i_t+n}+(1-p)]},$$

defined in a manner analogue to $p_t(n^j_t, n^i_t, s^j)$, being the probability that player j has obtained n^j_t draws of L by the end of period t, conditional on n^i_t and on the constraints of the stopping strategy s^j, including the one at the end of period t.[27]

Arguing as in Lemma 1, it can be shown that the distribution $p'_t(\cdot, n^i_t, s^j)$ first-order stochastically increases in n^i_t. Therefore, the probabilities $\hat{p}^H_t(n^i_t \,|\, s^j)$ and $\hat{p}^L_t(n^i_t \,|\, s^j)$ are respectively decreasing and increasing in n^i_t. In addition, $V^*_{t+1}(\cdot \,|\, s^j)$ is decreasing (from the induction hypothesis) and $V^*_{t+1}(n^i_{t+1} \,|\, s^j) \le (1/2)H$, for all $n^i_{t+1} \ge 0$. Hence, the payoff $U^*_{t+1}(n^i_t \,|\, s^j) = U_{t+1}[n^i_t \,|\, s^j, (N^i_\tau)_{\tau=t+1}^{T-1}]$ is decreasing in n^i_t.

At the end of period t, player i's expected gain from choosing to continue rather than to stop is

[27] In particular, $h'_t(n^j_t, s^j) \le \binom{t+1}{n^j_t}$ is the number of histories of player j consistent with player j having obtained n^j_t draws of L and the constraints of the stopping strategy s^j in periods $1, 2, \dots, t$. Notice that these constraints include the hypothesis that no draw of H has occurred.

$$\begin{aligned}
&\Delta V_t(n_t^i \,|\, s^j, (N_\tau^i)_{\tau=t+1}^{T-1}) \qquad \text{(A.5)}\\
&\quad = P[n_t^j < N_t^j \,|\, n_t^i, s^j]\,[\,U_{t+1}[n_t^i \,|\, s^j, (N_\tau^i)_{\tau=t+1}^{T-1}] - L\,]\\
&\qquad + P[n_t^j \geq N_t^j \,|\, n_t^i, s^j]\,(-L/2)\\
&\quad = P[n_t^j < N_t^j \,|\, n_t^i, s^j]\,[\,U_{t+1}[n_t^i \,|\, s^j, (N_\tau^i)_{\tau=t+1}^{T-1}] - L/2\,] - L/2
\end{aligned}$$

Using again the fact that an increase in n_t^i results in a stochastic dominant distribution for the unknown variable n_t^j, along with the fact that $U_{t+1}^*(\cdot \,|\, s^j)$ is decreasing, it follows that player i's gain $\Delta V_t[n_t^i \,|\, s^j, (N_\tau^i)_{\tau=t+1}^{T-1}]$ is decreasing in n_t^i, so that player i's best-response strategy in period t takes the form of a threshold rule, N_t^i.

Finally, since the probability $P[n_t^j < N_t^j \,|\, n_t^i, s^j]$ and the expected payoff functions $U_{t+1}(n_t^i \,|\, s^j, (N_\tau^i)_{\tau=t+1}^{T-1})$ are decreasing in n_t^i, it follows that the payoffs

$$\begin{aligned}
&V_t^c[n_t^i \,|\, s^j, (N_\tau^i)_{\tau=t+1}^{T-1}] = P[n_t^j < N_t^j \,|\, n_t^i, s^j]\, U_{t+1}[n_t^i \,|\, s^j, (N_\tau^i)_{\tau=t+1}^{T-1}],\\
&V_t^s[n_t^i \,|\, s^j, (N_\tau^i)_{\tau=t+1}^{T-1}] = (L/2) + P[n_t^j < N_t^j \,|\, n_t^i, s^j]\,(L/2)
\end{aligned}$$

and

$$\begin{aligned}
V_t^*(n_t^i \,|\, s^j) &= V_t[n_t^i \,|\, s^j, (N_\tau^i)_{\tau=t+1}^{T-1}]\\
&= \max\{\, V_t^c[n_t^i \,|\, s^j, (N_\tau^i)_{\tau=t+1}^{T-1}],\; V_t^s[n_t^i \,|\, s^j, (N_\tau^i)_{\tau=t+1}^{T-1}]\,\} \qquad \text{(A.6)}
\end{aligned}$$

are decreasing in n_t^i, completing the induction. □

Proof of Proposition 4. Similar to the proof of Lemma 3, we condition our continuation payoff calculations on player j having obtained no draw of H by the time of player i's decision.

In the continuation game starting at the end of period T, it is clear that the strategy profile in which each player stops immediately constitutes an equilibrium, independently of the players' strategies up to that period and associated beliefs.

In period $T-1$, suppose that the two players have followed symmetric strategies s' with stopping thresholds $\{N_t\}_{t=1}^{T-2}$ prior to that period; and that player j follows a threshold N_{T-1}^j in that period.[28] If player i has obtained $n_{T-1}^i > 0$ draws of L, then his expected gain from continuing to period T instead of stopping in period $T-1$ is given by equations (A.1) and (A.2) in the proof of Lemma 3.[29]

For parameters $H/L < (3-r-rq)/(2-rq)$, as argued in the proof of Lemma 3, we have $\Delta V_{T-1}(n_{T-1}^i, \,|\, n_{T-1}^j, s', N_{T-1}^j) < 0$, for all $n_{T-1}^i \geq 1$, $n_{T-1}^j \geq 0$, so that player i's continuation gain is $\Delta V_{T-1}(n_{T-1}^i \,|\, s', N_{T-1}^j) < 0$, for all $n_{T-1}^i \geq 1$. In this case, there is a unique equilibrium for the continuation game, with threshold $N_{T-1} = 1$.

[28] Notice that the players cannot observe one another's deviations, in particular, the deviation to continuing when a player's strategy prescribes stopping. Thus, in histories off the equilibrium path, a player's continuation or stopping payoff is not affected by his past behavior, it depends only on the number of L draws he has and the strategy of his opponent.

[29] Notice that player i's beliefs regarding the number of draws of his opponent, n_{T-1}^j, are independent of his opponent's continuation strategy, in particular, of the threshold N_{T-1}^j.

For parameters $H/L \geq (3 - r - rq)/(2 - rq)$, again as argued in the proof of Lemma 3, the payoff $\Delta V_{T-1}(n^i_{T-1} \,|\, s', N^j_{T-1})$ is decreasing in the number of draws n^i_{T-1}. In addition, under Condition SC, the payoffs $\Delta V_{T-1}(n^i_{T-1}, \,|\, n^j_{T-1}, s', N^j_{T-1})$ and, therefore, $\Delta V_{T-1}(n^i_{T-1} \,|\, s', N^j_{T-1})$ are increasing in player j's threshold N^j_{T-1}. Hence, the threshold characterizing player i's best-response strategy in period $T-1$, given by

$$BR^i_{T-1}(N^j_{T-1} \,|\, s') = \max\{n = 1, 2, \ldots, T-1 : \Delta V_{T-1}(n \,|\, s', N^j_{T-1}) > 0\} + 1,$$

with $BR^i_{T-1}(s^j) = 1$ when the set is empty, is an increasing function of the threshold N^j_{T-1} in the strategy s^j.[30]

The set $\{1, 2, ..., T\}$ is a lattice with respect to the order $\geq$, complete because of finiteness. Therefore, since the function $BR^i_{T-1}(\cdot \,|\, s')$ is increasing in the variable N^j_{T-1}, it has at least one fixed point. Hence, for each symmetric strategy $s' = \{N_t\}_{t=1}^{T-2}$ prior to period $T-1$, we can define the players' common threshold at time $T-1$ as the maximal fixed point of $BR^i_{T-1}(\cdot \,|\, s')$.

Moving backwards to periods $t = T-2, T-3, \ldots, 1$, suppose that for each symmetric strategy profile with stopping thresholds $\{N_\tau\}_{\tau=1}^{t}$ up to the end of period t, there is a symmetric equilibrium $s''[(N_\tau)_{\tau=1}^{t}]$ for the continuation game starting in period $t+1$, with thresholds that depend on $\{N_\tau\}_{\tau=1}^{t}$ (induction hypothesis).

Suppose that the two players have followed a symmetric threshold strategy s' up to the end of period $t-1$. We need to show that there is a threshold N in period t such that the continuation strategy $(N, s''(s', N))$ forms a symmetric equilibrium for the continuation game starting in period t, where $s''(s', N)$ is the symmetric equilibrium provided by the induction hypothesis for the continuation game starting in period $t+1$, when the players have followed strategies (s', N) up to the end of period t.

We first show that each player's best-response threshold in period t is increasing in the corresponding threshold in his opponent's strategy, for any symmetric threshold strategy s' the players have followed up to the end of period $t-1$ and for a symmetric continuation strategy determined by the induction hypothesis.[31] Subsequently, since the set of all thresholds in period t forms a finite lattice, we invoke a fixed-point theorem to conclude that the players' best-response function has a fixed point N, determining a symmetric equilibrium for the continuation game at time t.

Consider a change, first, of the players' symmetric threshold in period t from N to $N+1$, and second, of the players' symmetric equilibrium strategy for the continuation game starting at $t+1$ from $s''(s', N)$ to $s''(s', N+1)$. We examine how this change affects a player's best response.

If player i has n^i_t draws of L, then his expected gain from continuing rather than stopping at the end of period t, against a strategy $s(s', M) = [s', M, s''(s', M)]$ of player j, is

$$\Delta V_t[(n^i_t \,|\, s(s', M)] = P(n^j_t \geq M \,|\, n^i_t, s')(-L/2)$$

[30] If $\tilde{N}^j_{T-1} > N^j_{T-1}$, then we have $\Delta V_{T-1}(n \,|\, s', \tilde{N}^j_{T-1}) > \Delta V_{T-1}(n \,|\, s', N^j_{T-1})$, for all $n = 1, 2, \ldots, T$, implying that $\{n \in \mathbb{N} : \Delta V_{T-1}(n \,|\, s', \tilde{N}^j_{T-1}) > 0\} \supseteq \{n \in \mathbb{N} : \Delta V_{T-1}(n \,|\, s', N^j_{T-1}) > 0\}$ and, therefore, that the best response is $BR^i_{T-1}(\tilde{N}^j_{T-1} \,|\, s') \geq BR^i_{T-1}(N^j_{T-1} \,|\, s')$, as required.

[31] In this argument, notice that we do not find a best-response for the entire continuation game; we only find each player i's best response in period t, in the game in which the player is restricted after period t to follow the strategy provided by the induction hypothesis, determined by player j's strategy. However, if this best responde is symmetric, as we eventually show, then it determines a best-response strategy (and because of symmetry, an equilibrium) for the entire continuation game.

$$+ P(n_t^j < M \,|\, n_t^i, s') \, [\, U_{t+1}[n_t^i \,|\, s(s', M)] - L \,],$$

where $U_{t+1}(n_t^i \,|\, s(s', M))$, defined recursively by equations (A.1)–(A.6) in the proof of Lemma 3, is player i's optimal expected payoff in the continuation game starting in period $t+1$, conditional on period $t+1$ being reached, with player j following a strategy $s(s', M)$. Since player j's continuation strategy $s''(s', M)$ is part of a symmetric equilibrium for that game, given (s', M), notice that the payoff $U_{t+1}(n_t^i \,|\, s(s', M))$ is achieved with player i also following the continuation strategy $s''(s', M)$.

When player i's conjecture about player j's strategy changes from $s(s', N))$ to $s(s', N+1)$, we have

$$\begin{aligned}\Delta V_t[n_t^i \,|\, s(s', N+1)] - \Delta V_t[n_t^i \,|\, s(s', N)] = {} & p_t(N, n_t^i, s') \, (-L/2) \\ & + P(n_t^j \le N \,|\, n_t^i, s') \, U_{t+1}[n_t^i \,|\, s(s', N+1)] \\ & - P(n_t^j \le N-1 \,|\, n_t^i, s') \, U_{t+1}[n_t^i \,|\, s(s', N)]\end{aligned}$$

Since player i cannot gain from deviating from $s''(s', N+1)$ to the strategy of surely stopping in period $t+1$, against $s''(s', N+1)$, in the continuation game following $(s', N+1)$, we have

$$\begin{aligned}& U_{t+1}[n_t^i \,|\, s(s', N+1)] \ge \\ & \quad \sum_{n_t^j=0}^{N} \frac{p_t(n_t^j, n_t^i, s')}{P(n_t^j \le N \,|\, n_t^i, s')} \, (1/2) \, \delta \, [\, p(n_t^i + n_t^j) \, (1 - (1 - rq)^2) \, (H - L) + L \,]\end{aligned}$$

In addition, in the continuation game following (s', N), we have

$$\begin{aligned}& U_{t+1}[n_t^i \,|\, s(s', N)] \le \\ & \quad \sum_{n_t^j=0}^{N-1} \frac{p_t(n_t^j, n_t^i, s')}{P(n_t^j \le N-1 \,|\, n_t^i, s')} \, (1/2) \, \delta \, [\, p(n_t^i + n_t^j) \, (1 - (1 - rq)^{2(T-t)}) \, (H - L) + L \,]\end{aligned}$$

that is, player i's optimal expected payoff cannot exceed what could be achieved if the two players shared L or H after performing maximal costless experimentation in the time remaining until final period T.

Therefore, after some rearrangement of the terms, we have

$$\begin{aligned}& \Delta V_t[n_t^i \,|\, s(s', N+1)] - \Delta V_t[n_t^i \,|\, s(s', N)] \ge \\ & \quad p_t(N, n_t^i, s') \, (1/2) \, [\, \delta \, p(n_t^i + N) \, (1 - (1 - rq)^2) \, (H - L) - (1 - \delta) \, L \,] \\ & \quad - \sum_{n_t^j=0}^{N-1} p_t(n_t^j, n_t^i, s') \, (1/2) \, \delta \, p(n_t^i + n_t^j) \, [\, (1 - rq)^2 - (1 - rq)^{2(T-t)} \,] \, (H - L)\end{aligned}$$

In addition, since the function $p(\cdot)$ is decreasing, we have

$$\begin{aligned}& \Delta V_t[n_t^i \,|\, s(s', N+1)] - \Delta V_t[n_t^i \,|\, s(s', N)] \ge \\ & \quad p_t(N, n_t^i, s') \, (1/2) \, [\, \delta \, p(2N) \, (1 - (1 - rq)^2) \, (H - L) - (1 - \delta) \, L \,] \\ & \quad - \sum_{n_t^j=0}^{N-1} p_t(n_t^j, n_t^i, s') \, (1/2) \, \delta \, p(n_t^j) \, [\, (1 - rq)^2 - (1 - rq)^{2(T-t)} \,] \, (H - L)\end{aligned}$$

Thus, for player i's expected gain from continuing at the end of period t to be

$$\Delta V_t[n_t^i \,|\, s(s', N+1)] \;\geq\; \Delta V_t[n_t^i \,|\, s(s', N)]$$

it is sufficient that

$$p_t(N, n_t^i, s') \left[p(2N)\,[1-(1-rq)^2] \;-\; \tfrac{1-\delta}{\delta}\, \tfrac{L}{H-L} \right]$$
$$+ \sum_{n_t^j=0}^{N-1} p_t(n_t^j, n_t^i, s')\; p(n_t^j)\,[(1-rq)^{2(T-t)} \;-\; (1-rq)^2] \geq 0$$

The expression on the left-hand-side is the expectation of a function increasing in n_t^j with respect to a distribution of n_t^j that is stochastically increasing in n_t^i, so it achieves its minimal value for $n_t^i = 1$. Hence, the above inequality follows directly from Condition SC.

Hence, under Condition SC, for each strategy s' prior to period t, for each n_t^i, player i's expected gain $\Delta V_t[n_t^i \,|\, s(s', N_t^j)]$ from continuing instead of stopping at the end of period t is increasing in the threshold N_t^j parameterizing player j's continuation strategy $s''(s', N_t^j)$. Thus, for each strategy s' prior to period t, the threshold N_t^i parameterizing player i's best-response continuation strategy $s''(s', N_t^i))$ in period t,

$$BR_t^i(N_t^j \,|\, s') \;=\; \max\{n = 1, 2, \ldots, t :\; \Delta V_t[n \,|\, s(s', N_t^j)] \;>\; 0\} \;+\; 1,$$

with $BR_t^i(N_t^j \,|\, s') = 1$ when the set is empty, is an increasing function of the threshold N_t^j in player j' strategy $[s', N_t^j, s''(s', N_t^j)]$.

The set $\{1, 2, ..., t+1\}$ of possible thresholds in period t is a lattice with respect to the order $\geq$, complete because of finiteness. Therefore, since the function $BR_t^i(\cdot \,|\, s')$ is increasing in N_t^j, it has at least one fixed point.

For each symmetric threshold strategy s' prior to period t, we define the players' common threshold N_t at period t as the maximal fixed point of $BR_{T-1}^i(\cdot \,|\, s')$; and by construction, the continuation strategy $(N_t, s''(s', N_t))$ forms a symmetric equilibrium for the game starting at period t, when the two players have the beliefs induced by the strategy s' that they have followed prior to that period.

The argument concludes when it defines a threshold N_1 for the first period of the game, with the impled strategy $[N_1, s''(N_1)]$ forming a symmetric perfect Bayesian equilibrium for the entire game. □

Proof of Proposition 5. In the case of public learning, suppose that player i has obtained $n_1^i = 1$ draw of L in period $t = 1$ and faces an opponent who will continue to period $T = 2$, the last period of the game. If player j has obtained $n_1^j = 1$ draw of L in period $t = 1$, then player i's expected payoff from continuing to period $T = 2$ is

$$v_1(1,1) \;=\; \delta\,[\,L/2 \;+\; p^H(2)\,(H-L)/2\,]$$

If player j has obtained $n_1^j = 0$ draw of L in period $t = 1$, then player i's expected payoff from continuing to period $T = 1$ is

$$v_1(1,0) \;=\; \delta\,[\,L/2 \;+\; p^H(1)(H-L)/2 \;+\; (1-r)\,[1-p(1)\,rq\,]\,(L/2)\,]$$

Since all terms are positive and $p^H(1) > p^H(2)$, it follows that $v_1(1,0) > v_1(1,1)$.

Using some simple algebraic manipulations, it is easy to check that the inequalities $v_1(1,1) \geq L$ and $v_1(0,1) < L$ are equivalent respectively to conditions (8) and (9).

Now, consider the strategy in which a player continues at the end of period $t = 1$, independently of the number of draws he and his opponent have. For this strategy to be part of a symmetric equilibrium, it is necessary and sufficient that $v_1(1,0) \geq L$ and $v_1(1,1) \geq L$, a condition that reduces to $v_1(1,1) \geq L$, which is equivalent to condition (8).

Similarly, consider the strategy in which a player continues at the end of period $t = 1$ if and only if he has received no draw. For this strategy to be part of a symmetric equilibrium, it is necessary and sufficient that $v_1(1,1) < L$ and $v_1(1,0) < L$, a condition that reduces to $v_1(0,1) < L$, which is equivalent to condition (9).

Finally, consider the strategy in which a player continues at the end of period $t = 1$ if and only if either he or his opponent has failed to obtain a draw. For this strategy to be part of a symmetric equilibrium, it is necessary and sufficient that $v_1(1,1) < L$ and that $v_1(1,0) \geq L$, i.e., that conditions (8) and (9) both fail.

Looking at the corresponding setting under private learning, when condition (7) holds, by Proposition 4, there must exist at least one Bayesian equilibrium.

Suppose that player i has obtained one draw of L in period $t = 1$ and faces an opponent who will continue to period $T = 2$ unless he obtains H. Then player i's expected payoff from continuing (and stopping) at $T = 2$, conditional on his opponent having not obtained H, is

$$v_1(1) \;=\; p_1(0,1)\, v_1(1,0) \;+\; [1 - p_1(0,1)]\, v_1(1,1)$$

Using the expressions for $v_1(1,0)$ and $v_1(1,1)$ and applying some simple algebraic manipulations, it is easy to show that the inequality $v_1(1) < L$ is equivalent to condition (10).

When $v_1(1) \geq L$, the strategy profile in which each player continues to period $T = 2$ unless he obtains H forms a symmetric equilibrium under private learning.

When $v_1(1) < L$, this strategy profile is no longer an equilibrium. In this case, player i's expected payoff from continuing with one draw of L to period $T = 2$ against an opponent who will stop as soon as he receives one draw is

$$u_1(1) \;=\; p_1(0,1)\, v_1(1,0)$$

Therefore, the strategy profile in which each player stops if he obtains a draw at $t = 1$ forms a symmetric equilibrium under private learning if and only if

$$u_1(1) \;<\; p_1(0,1)\, L \;+\; [1 - p_1(0,1)]\,(L/2)$$

which is true when conditions (7) and (10) hold.

We conclude the proof by comparing the equilibria under public and private learning.

Under condition (8), we have that $v_1(1,0) \geq L$ and $v_1(1,1) \geq L$, so that $v_1(1) \geq L$. Therefore, in both settings, the two players continue to period $T = 2$ unless they obtain H and then stop, for the same equilibrium outcomes.

Similarly, under condition (9), we have $v_1(1,0) < L$ and $v_1(1,1) < L$, so that $v_1(1) < L$. In both settings, each player stops either as soon as he obtains a draw, again for the same equilibrium outcomes.

Finally, if conditions (8) and (9) both fail, we have $v_1(1,1) < L$ and $v_1(1,0) \geq L$, so, under public learning the two players stop at $t = 1$ if and only if they both obtain draws. Under private learning, when condition (10) holds, the game will stop in period $t = 1$ even with a single draw, for a shorter expected experimentation horizon. On the other hand, when condition (10) fails, the game will continue to period $t = T = 2$ unless H is obtained, for a longer expected experimentation horizon.

When condition (10) holds, less experimentation under private learning implies also lower expected payoffs, since the generated welfare is respectively increasing / decreasing in N, the total number of L draws that the players obtain by the time they stop experimenting. □

Proof of Proposition 6. For any probability parameters $r, q \in [0, 1]$, we need to show that if condition (10) is satisfied for some probability $p \in [0, 1]$, then it is also satisfied for all probabilities $p' \leq p$. For this, we need that the LHS in inequality (10) is increasing in p. Equivalently, we show that the continuation payoff $v_1(1) = v_1(1; p)$, defined in the proof of Proposition 5, is increasing in p.

Suppose first that $H/L > 3/2$. Then the continuation payoffs $v_1(1, 1; p)$ and $v_1(1, 0; p)$ are both increasing in p, with $v_1(1, 0; p) \geq v_1(1, 1; p)$, for all $p \in [0, 1]$. In addition, the beliefs $p_1(0, 1) = p_1(0, 1; p)$ are increasing in p. Therefore,

$$\begin{aligned} \tfrac{\partial}{\partial p} v_1(1; p) = {} & \tfrac{\partial}{\partial p} p_1(0, 1; p)\, [v_1(1, 0; p) - v_1(1, 1; p)] \\ & + p_1(0, 1; p)\, \tfrac{\partial}{\partial p} v_1(1, 0; p) \;+\; [1 - p_1(0, 1; p)]\, \tfrac{\partial}{\partial p} v_1(1, 1; p) \quad > \; 0, \end{aligned}$$

since all terms are positive, so that $v_1(1; p)$ is increasing in p.

Finally, when $H/L < 3/2$, then

$$\begin{aligned} & p(1)\,[1 - (1 - rq)^2]\,(H - L)/L \;+\; (1 - r)[1 - p(1)\,rq] \\ & \quad < p(1)\,[1 - (1 - rq)^2]\,, (1/2) \;+\; (1 - r)[1 - p(1)\,rq] \\ & \quad = 1 \;-\; r\,[1 - p(1)\,rq\,(1 - q/2)] < 1 < (2 - \delta)/\delta, \end{aligned}$$

so that condition (9) and therefore condition (10) are satisfied for all probabilities $p \in [0, 1]$, for the result to hold trivially. □

Proof of Proposition 7. In the case of public learning, by Proposition 1, the equilibrium is characterized by stopping thresholds N_1 and N_2 on the number of L draws that the two players obtain, respectively for the case in which both players or only a single player receives these draws. Since conditions (8) and (9) do not hold,[32] it follows from arguments similar to those used in Proposition 5 that $N_1 = 2$ while $N_2 \geq 2$.

In the case of private learning, we construct the unique equilibrium of the game by arguing backwards, looking at the continuation games in periods $t = T - 1, \ldots, 1$. As we will show, since condition (10) holds, in equilibrium, each player stops as soon as he receives a draw.

In period $t = T - 1$, suppose that each player i has obtained $n^i_{T-1} \leq 1$ draws of L and believes with certainty that $n^j_{T-2} = 0$. In this continuation game, the players' problem is identical to that analyzed in Proposition 5. Thus, each player i's decision to continue or to stop in period $T - 1$, when $n^i_{T-1} = 1$, depends on condition (10). Since this condition holds, as argued in the proof of Proposition 5, player i is better-off stopping with one draw, even if his opponent will not stop in the current period unless he obtains a draw of H. Therefore, in this continuation game, there is a unique equilibrium, with each player i stopping if he has a draw.

The above argument also applies to continuation games starting in period $T - 1$, in which either $n^i_{T-1} > 1$ or player i attaches positive probability to $n^j_{T-2} > 0$, or both. Again, since

[32] That condition (8) does not hold is implied by condition (10).

condition (10) holds and since player i' beliefs about the possibility of H are more pessimistic than those for $n^i_{T-1} = 1$, each player i will prefer to claim L in the current period, even if he knows that player j will stop only if he obtains H. Therefore, in all continuation games starting in period $T - 1$, for any beliefs of each player i regarding n^j_{T-2}, there is a unique continuation equilibrium, characterized by a threshold $N_{T-1} = 1$.

Moving to $t = T - 2$, in any continuation game starting in that period, player i will trivially continue if $n^i_{T-2} = 0$. Furthermore, if $n^i_{T-2} \geq 1$, player i knows that even if he continues, he will surely stop in period $T - 1$. Thus, he faces a two-period problem identical to that the continuation game starting at $T - 1$. It follows that there is a unique continuation equilibrium, with threshold $N_{T-2} = 1$.

Moving backwards, replicating the above argument, we conclude that under private learning there is a unique equilibrium, characterized by thresholds $N_t = 1$, for all $t = 1, \dots, T - 1$.

Finally, the comparison of public and private learning in terms of expected experimentation length and payoffs is trivial. For any sequence of draws that the two players may receive, if the players stop under public learning, then they also stop under private learning. And for some sequences, for example, a sequence involving exactly one draw at $t = 1$, experimentation will stop under private learning but will continue for at least one more period under public learning.[33] □

Appendix B. The two-period problem under private learning

In this appendix, we identify all symmetric equilibria in the two-period problem under private learning. In particular, we describe mixed-strategy equilibria.

Since a player is better-off stopping if he draws H and continuing if he receives no draw in period $t = 1$, the investigation of each player's incentives reduces to determining his best response when he has received a draw of L in the first period.[34]

Suppose that a player's opponent continues to period $t = 2$ if he has a draw of L. Then that player is better-off also continuing to $t = 2$ if and only if

$$p_1(0,1)\, v_1(1,0) \;+\; [1 - p_1(0,1)]\, v_1(1,1) \;\geq\; L,$$

where the terms

$$\begin{aligned} v_1(1,1) &= \delta\, [L/2 \;+\; p^H(2)\,(H - L)/2], \\ v_1(1,0) &= \delta\, [L/2 \;+\; p^H(1)(H - L)/2 \;+\; (1 - r)\,[1 - p(1)\, rq], \end{aligned}$$

defined in the proof of Proposition 5, express the player's expected payoff when he continues to $t = 2$, conditional on his opponent continuing respectively with one and with no draw of L. A simple algebraic manipulation shows that this equivalent to

$$\begin{aligned} &[p_1(0,1)\, p(1) + (1 - p_1(0,1))\, p(2)]\, [1 - (1 - rq)^2]\, \frac{H - L}{L} \\ &\qquad + \; p_1(0,1)\, (1 - r)\, [1 - p(1)\, rq] \geq \frac{2 - \delta}{\delta} \end{aligned} \tag{B.1}$$

[33] That more experimentation results in this case in higher expected payoff for the players follows from the fact that the cooperative optimal threshold is higher than that under public learning.

[34] As in the rest of the paper, we ignore the trivial equilibrium in which each player stops at the end of the first period, independently of his draw in it.

Suppose that a player's opponent stops in period $t = 1$ if he has a draw of L. Then that player is better-off also stopping if and only if

$$p_1(0,1)\, v_1(1,0) \;\le\; p_1(0,1)\, L \;+\; [1 - p_1(0,1)]\,(L/2))$$

This is equivalent to the condition

$$p_1(0,1)\;\delta\,[\,L \,+\, p(1)\,(1-(1-rq)^2)\,(H-L) + (1-r)(1-p(1)rq)\,L)\,] \\ \le\; [1 \,+\, p_1(0,1)]\, L \qquad \text{(B.2)}$$

The two conditions describing a player's incentives to continue or to stop at the end of period $t = 1$ against an opponent following the same strategy lead to the following result:

Proposition 8. *In the two-period problem under private learning, depending on whether the parameters of the problem satisfy the conditions* (B.1) *and* (B.2)*, the following symmetric equilibria occur:*

a. *When* (B.1) *holds and* (B.2) *fails, there is a single equilibrium, in pure strategies, with each player continuing to period* $t = 2$ *with a draw of* L*.*
b. *When* (B.1) *fails and* (B.2) *holds, there is a single equilibrium, in pure strategies, with each player stopping in period* $t = 1$ *with a draw of* L*.*
c. *When both* (B.1) *and* (B.2) *hold, there are two equilibria in pure strategies, described in (a) and (b). In addition, there is a mixed-strategy equilibrium.*
d. *When both* (B.1) *and* (B.2) *fail, there is a single equilibrium, in mixed strategies.*

Proof. To prove the above result for the pure-strategy equilibria, it is straightforward to check that the conditions that hold or fail in each case respectively establish the best-response behavior required for the equilibria claimed in that case or violate the equilibrium requirements for the remaining symmetric strategy profiles.

So, it suffices to examine when a mixed-strategy equilibrium exists. For this, suppose that a player's opponent follows a strategy such that if he has a draw of L in period $t = 1$, he continues to the next period with probability $\alpha \in [0, 1]$. For the player to be indifferent between continuing and stopping at the end of period $t = 1$, when he has one draw of L, we must have

$$p_1(0,1)\, v_1(1,0) \,+\, [1-p_1(0,1)]\,\alpha\, v_1(1,1) \;=\; p_1(0,1)\, L \,+\, [1-p_1(0,1)]\,(1+\alpha)\,(L/2)$$

Rearranging the terms gives the equation

$$p_1(0,1)\,[v_1(1,0) - L] + [1 - p_1(0,1)]\,\alpha\,[v_1(1,1) - L] - [1 - p_1(0,1)]\,(1-\alpha)\,(L/2) \;=\; 0,$$

which can be also expressed as

$$[1 - p_1(0,1)]\,\alpha\,[v_1(1,1) - (L/2)] \,+\, p_1(0,1)\,[v_1(1,0) - (L/2)] \,-\, (L/2) \;=\; 0,$$

so that

$$\alpha^* \;=\; \frac{(L/2) - p_1(0,1)\,[v_1(1,0) - (L/2)]}{[1 - p_1(0,1)]\,[v_1(1,1) - (L/2)]}$$

When conditions (B.1) and (B.2) both hold, by substituting (B.2) into (B.1), we can show that the sign of the denominator in the expression for α^* is positive.[35] Thus, condition (B.1) implies

[35] Notice that this means that condition (7) in section 6, i.e., the sufficient condition for existence of equilibrium in pure strategies, holds.

that $\alpha^* \leq 1$ while condition (B.2) implies that $\alpha^* \geq 0$, for an equilibrium in mixed strategies to exist.[36]

Vice versa, when conditions (B.1) and (B.2) both fail, the same argument with reverse inequalities shows that the sign of the denominator in the expression for α^* is negative. Therefore, the failure of condition (B.1) implies that $\alpha^* \leq 1$ and similarly the failure of condition (B.2) implies that $\alpha^* \geq 0$.

Finally, when one of the conditions (B.1) and (B.2) holds while the other condition fails, for either a positive or negative denominator in the expression for α^*, it is easy to show that either $\alpha^* > 1$ or $\alpha^* < 0$, so that no mixed-strategy equilibrium exists. □

References

Akcigit, U., Liu, Q., 2015. The role of information in innovation and competition. J. Eur. Econ. Assoc. 14, 828–870.
Arrow, K.J., 1969. Classification notes on the production and transmission of technological knowledge. Am. Econ. Rev. 59, 29–35.
Barwich, A.-S., 2019. The value of failure in science: the story of grandmother cells in neuroscience. Front. Neurosci. 13, 1121. https://doi.org/10.3389/fnins.2019.01121.
Bobtcheff, C., Bolte, J., Mariotti, T., 2017. Researcher's dilemma. Rev. Econ. Stud. 84, 969–1014.
Bobtcheff, C., Levy, R., Marriotti, T., 2021. Negative results in science: Blessing or (winner's) curse? Working Paper.
Bornmann, L., Ye, A., Ye, F., 2018. Identifying landmark publications in the long run using field-normalized citation data. J. Doc. 74, 278–288.
Boyarchenko, S., Levendorskii, S., 2014. Preemption games and Lévy uncertainty. Games Econ. Behav. 88, 354–380.
Broad, W.J., 1981. The publishing game: getting more for less. Science 211, 1137–1139.
Dong, M., 2021. Strategic experimentation with asymmetric information. Working paper.
Fang, F.C., Casadevall, A., 2012. Intense competition among scientists has gotten out of hand. Sci. Am. 307 (2), 13–17.
Gaston, J., 1973. Originality and Competition in Science. The University of Chicago Press, Chicago.
Hagstrom, W.O., 1974. Competition in science. Am. Sociol. Rev. 39, 1–18.
Halac, M., Kartik, N., Liu, Q., 2017. Contests for experimentation. J. Polit. Econ. 125, 1523–1569.
Heidhues, P., Rady, S., Strack, P., 2015. Strategic experimentation with private payoffs. J. Econ. Theory 159, 531–551.
Hopenhayn, H.A., Squintani, F., 2011. Preemption games with private information. Rev. Econ. Stud. 78, 667–692.
Hoppe, H.C., Lehmann-Grube, U., 2005. Innovation timing games: a general framework with applications. J. Econ. Theory 121, 30–50.
Keller, G., Rady, S., Cripps, M., 2005. Strategic experimentation with exponential bandits. Econometrica 73, 39–68.
Lawrence, P.A., 2016. The last 50 years: mismeasurement and mismanagement are impeding scientific research. Curr. Top. Dev. Biol. 116, 617–631.
Lippman, S.A., Mamer, J.W., 1993. Preemptive innovation. J. Econ. Theory 61, 104–119.
Margaria, C., 2020. Learning and payoff externalities in an investment game. Games Econ. Behav. 119, 234–250.
Milgrom, P., Shannon, C., 1994. Monotone comparative statics. J. Econ. Theory 61, 104–119.
Mohs, R.C., Greig, N.H., 2017. Drug discovery and development: role of basic biological research. Alzheimers Dement (N Y) 3, 651–657.
Moscarini, G., Squintani, F., 2010. Competitive experimentation with private information: the survivor's curse. J. Econ. Theory 145, 639–660.
Nature Publishing Group, 2012. Must try harder. Nature, Editorial 483, 509.
Newman, M.E.J., 2009. The first-mover advantage in scientific publication. Europhys. Lett. 86, 68001.
Parkes, E., 2019. Scientific progress is built on failure. Nature, Comments and Opinion. https://doi.org/10.1038/d41586-019-00107-y.
Rosenberg, D., Salomon, A., Vieille, N., 2013. On games of strategic experimentation. Games Econ. Behav. 82, 31–51.
Sabatier, M., Chollet, B., 2017. Is there a first mover advantage in science? Pioneering behavior and scientific production in nanotechnology. Res. Policy 46, 522–533.

[36] To be precise, for the mixed-strategy equilibrium to exist, conditions (B.1) and (B.2) should both hold as strict inequalities.

Spatt, C.S., Sterbenz, F.P., 1985. Learning, preemption, and the degree of rivalry. Rand J. Econ. 16, 84–92.
Stein, C., Hill, R., 2021. Race to the bottom: how competition to publish first can hurt scientific quality. SIEPR Policy Brief, 1–5.
Wagner, P.A., Klein, N., 2022. Strategic investment and learning with private information. J. Econ. Theory 204.