

Social Experimentation with Interdependent and Expanding Technologies

Umberto Garfagnini
ITAM School of Business

Bruno Strulovici
Northwestern University

June 30, 2015

Abstract

How do successive, forward-looking agents experiment with interdependent and endogenous technologies? In this paper, trying a radically new technology not only is informative of the value of similar technologies, but also reduces the cost of experimenting with them, in effect expanding the space of affordable technologies. Successful radical experimentation has mixed effects: it improves the immediate outlook for further experimentation but decreases the value and the *marginal* value of experimentation in a longer term, resulting in less ambitious ‘incremental’ experimentation and in a reduced size of radical experimentation. Incremental experimentation lowers the option value of similar technologies, which may spur a new wave of radical experimentation. However, experimentation eventually stagnates for all parameters of the model.
JEL Codes: C73, D83, O3

1 Introduction

From the ubiquitous use of transistors and integrated circuits in the electronics industry to the exploitation of chromatography and DNA sequencing in medical research, contemporary firms may be seen as building on the innovation of their predecessors: discoveries that took place decades or years ago are integrated and combined by engineers and scientists to serve as a stepping stone for new innovation (Baumol, Litan, and Schramm 2007). This process has a number of features best described as a sequence of experimentations by successive agents. Learning occurs through two channels: the transmission of information across generations of agents, and the experimentation that each agent performs during his lifespan.

This type of experimentation cannot be framed as a standard multi-armed bandit problem. Firstly, the “arms,” or technologies, have interdependent values. For example, the emergence

We are grateful to Alessandro Bonatti, Steve Callander, Xavier Durand, Michal Fabinger, Michael Mandler, Philip Marx, Kiminori Matsuyama, Joel Mokyr, Francisco Ruiz-Aliseda, Simon Wilkie, Huanxing Yang, and seminar participants at the Spring 2011 Midwest Theory Conference, Banco de México, and the 2012 North American Winter Meeting of the Econometric Society. Strulovici gratefully acknowledges financial support from the NSF under Grant No. 1151410 and the Alfred P. Sloan Foundation. Email addresses: umberto.garfagnini@itam.mx and b-strulovici@northwestern.edu.

of a new drug offers a test not only for the value of that drug, but also for the value of drugs based on similar chemical compounds or on the same treatment strategy, as illustrated by the treatment of various cancers.¹ Secondly, technologies that seemed prohibitively costly to explore decades ago are now within reach. The domain of affordable technologies is thus expanding over time, and this happens through particular innovations which are “radical,” i.e., going beyond the mere combination of previous innovations. Radical innovations require substantive investments which lay the foundation for subsequent “incremental” improvements.² Finally, to analyze the long-run dynamics and potential stagnation of experimentation, a satisfactory model must allow technological and payoff sets which cannot be exhausted in finite time, which means that these sets must be unbounded.

Formulating a tractable model of interdependent and unbounded technologies and payoffs which captures an option value of experimentation is challenging. In particular, it is well-known that the standard bandit model increases substantially in complexity once arms become correlated. The classic index result of Gittins (1979) no longer holds. As we will see, experimentation with interdependent technologies has the implication that experimentation with one technology affects the option value of experimenting with other related technologies.

This paper provides a novel framework which addresses these issues. At each period, a new forward-looking firm, or agent, is born and lives for two periods.³ The agent observes the technologies chosen by previous agents and their outcome. In our model, technologies lie on the positive real half-line and their associated (but unknown) outcomes are drawn according to the path of a Brownian motion. Technologies are thus interdependent: by trying a technology z between two known technologies x and y , an agent reduces the variance and updates the expected value of all technologies between x and y . This feature of the model allows us to capture “incremental” experimentation as a convex combination of explored technologies.

The agent can alternatively choose to perform radical experimentation, such as developing an entirely new treatment to cure a disease, or a new procedure to produce a good, such as nuclear power for commercial use in the sixties and seventies.⁴ This possibility is represented by the

¹For example, Novartis initially developed Afinitor to treat renal cell carcinoma. Due to Afinitor’s success, scientists at Novartis believed the drug could be used to treat other types of cancer. They began testing for breast cancer and later received approval. See “*The CEO of Novartis: On Growing After a Patent Cliff*” in the *Harvard Business Review* (December 2012).

²A natural example is the birth of the internet, which may be traced back to the “radical” thinking of J.C.R. Licklider, a scientist at the Advanced Research Projects Agency, who first conceived the idea of a network of computers which led to the creation of ARPANET. See Ruttan (2006, p. 116).

³The option value of experimentation, which drives incentives to experiment, would not arise with myopic or one-period agents. We consider agents living for more than two periods in Section 7.

⁴Worldwide nuclear capacity grew from 1 gigawatt to 100 gigawatts in these two decades. The number of active reactors grew similarly during that period. See the annual publication of “*Nuclear Power Reactors in the World*”

choice of technologies which lie beyond the rightmost known technology x . These technologies have a variance which increases linearly in the distance from x thus generating a higher option value of experimentation compared to incremental experimentation. Such radical experimentation, however, entails a specific cost of fundamental research, which increases with the novelty, or size, of experimentation.

The main research questions addressed here concern the type and size of experimentation as it unfolds over time: Does successful experimentation encourage further radical experimentation, or does it spur a wave of less ambitious, incremental experimentation? How does the size of radical experimentation depend on past outcomes? Does radical experimentation persist in the long-run? Or does it converge to a well-defined technological standard, despite the unboundedness of the technological and payoff spaces?

These questions may be tractably studied thanks to the Brownian model, which captures a continuum of correlated (deterministic) bandits with unbounded payoffs. The Markovian structure of Brownian motion allows us to establish a value for each interval between previously tried technologies, instead of the classic approach of assigning a value for each individual technology. We develop new techniques to establish properties of this value and use them to characterize the short and long run dynamics of experimentation.

Successful past experimentation turns out to have two opposite effects on subsequent experimentation. Firstly, past successes reduce not only the *value* of radical experimentation, but also its *marginal value*. Intuitively, if past technologies provide high utility, there is little immediate need to experiment further. More precisely, it becomes less likely that these technologies will be surpassed by radical experimentation. Likewise, the added benefit from performing more radical experimentation decreases with past successes, as those successes reduce the probability that the option value created by this experimentation will be exercised. Formally, the value of radical experimentation is *submodular* in experimentation size and the value of past technologies. Secondly, however, a failed radical experimentation reduces the prospects of building on the corresponding technology. Suppose that in a given field what was the most promising and groundbreaking approach turns out to be a disappointment. Such a failure can create much pessimism about further new research in that area and result in less ambitious or 'fundamental' research.

The combination of these two effects creates a complex dynamic of experimentation: Highly successful radical experimentation immediately begets further radical experimentation. However, should that new attempt be less successful, the initial success becomes a drag on radical experimentation: it reduces the likelihood of further radical experimentation, and the size of such experimentation.

published by the International Atomic Energy Agency.

The dynamics of experimentation may be more fully described as follows. When incremental experimentation takes place, its informational option value decreases. This makes radical experimentation more attractive and may spur a succession of radical experimentations. As new technologies become available, however, the option value of incremental experimentation gets replenished, making it more attractive, and radical experimentation gives again way to incremental experimentation, generating a cycle of sort between the two types of experimentation.

Eventually, experimentation stagnates with probability one, for all specifications of the model. The intuition may be summarized as follows: Suppose that recent radical experimentation has been extremely disappointing relative to known technologies, which must eventually happen for any path. Then, because such radical experimentation serves as the basis of further radical experimentation, agents prefer to stay within the confines of incremental experimentation. Therefore, an informational cascade of sort arises, in which no agent accepts to bear the cost of radical experimentation, and the outlook for further radical experimentation is henceforth frozen at a negative level. Thereafter, only incremental experimentation occurs, and technological knowledge converges to a well-defined limit. Establishing this convergence is particularly challenging because both the set of technologies and the range of their payoffs are unbounded – as they should be in any model investigating whether radical experimentation is sustainable in the long run.

This work draws on several strands of literatures, concerned with experimentation, social learning, and innovation. The literature on multi-armed bandits and strategic experimentation typically considers a fixed set of independent arms. Here, in contrast, technologies are modeled as a continuum of “arms” with interdependent payoffs and varying costs. Moreover, the cost of experimenting with a given technology depends on the technologies that have been tried in the past. These features are new, to our knowledge, in the experimentation literature.⁵

Our baseline model focuses on a succession of agents living for two periods. This assumption is made for several reasons. Firstly, it makes the model tractable. It is well-known, and is easy to show, that when arms are correlated the optimal experimentation policy does not follow a Gittins-like index policy. An example of this is provided in Appendix E. Secondly, the structure of the model brings it closer to the social learning literature which, since Banerjee (1992) and Bikhchandani, Hirshleifer, and Welch (1992), has been an important source of inspiration to study learning dynamics, and has served as the main paradigm to analyze knowledge accumulation by successive agents from the adoption of new technologies to cultural change. These models examine the asymptotic efficiency and fragility of social learning under restrictive assumptions (myopic agents, exogenous signals, stationary and limited choice sets), however, which make them ill-suited

⁵Callander (2011) and Jovanovic and Rob (1990) also consider Brownian uncertainty, but their agents are myopic. The relation to those papers is detailed below.

to study important knowledge accumulation processes, such as technological innovation.⁶ It also makes social learning models unsuitable to study whether radical experimentation is self-sustainable in the long run or requires external intervention, a question that has received longstanding interest from economists and policymakers alike.⁷

Our model builds on Callander (2011), who uses Brownian motion to model the uncertainty of policy values. In that model, however, agents are myopic and their goal is to find a zero of some Brownian path. In particular, there is no uncertainty about the value of the optimal policy. The combination of forward-looking agents and uncertainty in the value of the optimal policy creates arguably richer dynamics. Experimentation cycles and stagnation, absent from Callander (2011), are central features of our analysis.

Jovanovic and Rob (1990) consider a countable infinity of research dimensions, each modeled as an independently distributed standard Brownian path over a fixed interval. Interpreting each new dimension as a separate technology, it is unclear why the outcomes of these technologies should be completely independent of one another. In reality, even radical breakthroughs build on existing technologies. It is also counterintuitive, in that model, that a breakthrough cannot be followed up by a further push in the same direction. Lining up technologies on a single dimension allows us to investigate the consequences of path dependence, which is an important empirical phenomenon in the history of innovation. The independent dimensions assumption also makes impossible to study how the size of radical experimentation is affected by past successes, a key aspect of our analysis. The mechanisms underlying experimentation (or innovation) cycles are also different: In Jovanovic and Rob, agents continue to pursue radical experimentation after disappointing outcomes from previous radical experimentation, owing to the independence of Brownian processes driving each new technology. Here, by contrast, successful radical experimentation begets more radical experimentation because path dependence affects agent's beliefs.

Both Callander (2011) and Jovanovic and Rob (1990) restrict attention to myopic agents. The option value of experimentation, at the heart of the experimentation literature and of the present paper, is thus absent from those works. Jovanovic and Rob circumvent this issue by allowing the agent to learn, at some fixed cost, the value of a new “technique,” and then decide whether to try the new technique or use an old one. By disentangling information acquisition from technological choice, that model restores some learning from otherwise myopic agents. In many applications, technological choices require irreversible investments so that search and experimentation cannot be thus disentangled.

Finally, our model provides a stark prediction for long-run experimentation, with the emergence

⁶Mueller-Frank and Pai (2013) study social learning when agents' signals are costly and endogenously acquired.

⁷See, e.g., Griliches (1992), Hall (1996), Gordon (2012). Recent press coverage includes Slywotsky (2009), Zakaria (2011). Policymakers' interest was highlighted by President Obama's State of the Union Address in 2011.

of a technological standard after finitely many cycles of radical and incremental experimentation. The analysis establishes that experimentation and beliefs converge to a well-defined limit despite the fact that both the technological and payoff spaces are non compact. Showing this result is challenging. One technical contribution of the paper is to extend to unbounded sets (both for the *domain* of technologies and the *range* of their payoffs) the techniques developed by Easley and Kiefer (1988) to show the existence of a well-defined limit for agents’ beliefs resulting from experimentation.

After describing the baseline model in Section 2, we analyze incentives for radical experimentation and introduce the value of radical experimentation in Section 3 and discuss incremental experimentation and the challenges presented by the correlation across technology payoffs in Section 4. Section 5 investigates the dynamics of social experimentation, including experimentation cycles and long-run stagnation. The social inefficiency of this stagnation is discussed in Section 6. Section 7 discusses how the insights developed in the baseline model can be generalized to agents living for more than two periods. Section 8 compares our mechanism of stagnation with those that have been presented in the literature on growth and discusses ways in which the baseline model could be modified to break our stagnation result. All the proofs are in the Appendix, which also contains an extension to the case in which the cost of radical experimentation depends on the outcome of past experimentation, and a discussion of patent policy.

2 A Model of Social Experimentation

An agent is born in each period $t \in \mathbb{N} = \{0, 1, \dots\}$, who lives for two periods, “young” and “old.” The agent is risk neutral and chooses at each period a technology x in the technological space $[0, \infty)$. The payoff $B(x)$ of technology x is initially unknown, except at the origin where $B(0) = 0$.⁸

A young agent inherits from the contemporary old agent the knowledge of all technologies and payoffs that have previously been tried. This information transmission is assumed to be costless and non-strategic and is the only interaction across those agents.⁹ The history \mathcal{H}_t at time t consists of all technology-payoff pairs that have been experienced in the past. We let \bar{B}_t denote the highest payoff among explored technologies, \hat{x}_t denote the rightmost explored technology, i.e., the current *frontier* of the explored domain, with associated payoff $\hat{B}_t \equiv B(\hat{x}_t)$, and $G_t \equiv \bar{B}_t - \hat{B}_t \geq 0$ denote the difference, or *gap*, between the payoffs of the best explored technology and the frontier technology.

⁸The function $B(\cdot)$ could represent utils instead of payoffs, as long as agents are expected-utility maximizers. One could also consider the larger domain \mathbb{R} . Radical experimentation would then have “left” and “right” components.

⁹The older generation is not affected by the choice of the younger one and has therefore no incentive to manipulate the transmission. It could benefit from selling the information. We discuss patents in Appendix F. In Niehaus (2011), intergenerational information transmission is costly, and altruistic older generations weight this cost of transmission with the potential benefit for the younger generation.

The payoff function B is distributed as a Brownian motion with drift κ and volatility $\sigma > 0$.¹⁰ The analysis assumes that $\kappa = 0$, which avoids complications with old-agent experimentation.¹¹ The payoff of a technology $x > \hat{x}_t$ has a normal distribution with mean \hat{B}_t and variance $\sigma^2(x - \hat{x}_t)$. Technologies to the right of \hat{x}_t thus have the same expected payoff, and a variance that increases with their distance from \hat{x}_t .

A technology x lying in some bounded interval $[x_l, x_r]$ with known endpoint payoffs $B(x_l)$ and $B(x_r)$ has a normally distributed payoff with mean

$$B(x_l) + \frac{B(x_r) - B(x_l)}{x_r - x_l}(x - x_l), \quad (1)$$

and variance

$$\frac{(x - x_l)(x_r - x)}{x_r - x_l}\sigma^2. \quad (2)$$

The expected payoff of a technology increases linearly from the endpoint technology with the worst payoff to the one with the highest payoff. The variance, instead, increases as one moves away from either endpoint, and is maximized at the midpoint of the interval. Owing to the Markovian structure of Brownian motion, observing the payoff of a technology in any given interval affects only the distribution of technologies lying in that interval.

Incremental vs. Radical Experimentation

The literature on growth and technological change has taken numerous approaches to define incremental and radical experimentation, which include informational and payoff components. The spatial structure of the model provides a natural way to dissociate those concepts. *Incremental experimentation* refers, in period t , to a technological choice $x \in [0, \hat{x}_t]$. This type of experimentation can be interpreted as a convex combination of, or marginal changes to old technologies. It involves technologies that are less risky to develop but also with more limited potential for success. *Radical experimentation* refers to a technology lying beyond the frontier \hat{x}_t . Radical experimentation is unbounded in its possibilities and it is characterized by higher variance which could lead to very successful technologies as well as complete failures.

Costs of Experimentation. Radical experimentation incurs a cost that depends on how far that technology lies from the current frontier. This cost represents the large initial investments that characterize fundamental research, together with its high uncertainty. In summary, the cost of experimentation is equal to $c(x - \hat{x})$, where c is twice continuously differentiable, increasing,

¹⁰More precisely, uncertainty is modeled with a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P})$ satisfying the usual regularity conditions (see, e.g., Karatzas and Shreve 1991) and whose outcomes are identified with the paths of a Brownian motion.

¹¹The case of an arbitrary drift is analyzed in the online appendix and does not overturn the stagnation theorem obtained with a zero drift.

convex, and such that i) $c(y) = 0$ for $y \leq 0$ and ii) either $c'(0) > 0$ or $c''(0) > 0$.¹² Incremental experimentation is assumed to be cheaper than radical experimentation and for simplicity we assume that it is costless. This assumption is, among others, justified by Romer (1990, p. 72): “Once the cost of creating a new set of instructions has been incurred, the instructions can be used over and over again at no additional cost. Developing new and better technologies is equivalent to incurring a fixed cost. This property is taken to be the defining characteristic of technology.” The qualitative results of the paper are robust to the introduction of a positive cost, provided that either i) exploitation of old technologies incurs the same cost, or ii) the cost of incremental experimentation vanishes as experimentation becomes arbitrarily close to known technologies.

This cost asymmetry is required to distinguish between incremental and radical experimentation, which is an important feature of the process of innovation in reality. The formulation also captures the idea that radical experimentation creates a positive *payoff* externality on future generations, in addition to an *informational* externality, because all technologies between the old and new frontiers become cheaper to explore after the initial investment has been made.

Each agent maximizes his total expected payoff, discounting his second period payoff by a factor $\delta \in (0, 1]$. An old agent has no value for information and thus always chooses the best explored technology (with zero drift, the expected payoff of any radical experimentation is equal to the payoff at the frontier). A young agent solves the optimization problem

$$U(\mathcal{H}_t) = \sup_{x \in [0, \infty)} E_{\mathcal{H}_t} [B(x) - c(x - \hat{x}_t) + \delta \max\{B(x), \bar{B}_t\}]. \quad (3)$$

3 Radical Experimentation

Incentives for radical experimentation are captured by the **value of radical experimentation**, $V^R(\mathcal{H}_t)$, which represents the highest lifetime payoff a ‘young’ agent can get when his first-period choice is restricted to radical experimentation, *in excess of the payoff guaranteed by exploitation*, \bar{B}_t . With this restriction on his first-period choice, a young agent’s value function is equal to $(1 + \delta) [\bar{B}_t + V^R(\mathcal{H}_t)]$.

Consider a history \mathcal{H}_t , with maximum explored payoff \bar{B}_t , frontier payoff \hat{B}_t , and gap $G_t = \bar{B}_t - \hat{B}_t$. Proposition 1 provides a useful representation for the value of radical experimentation.

PROPOSITION 1 (VALUE OF RADICAL EXPERIMENTATION) *(i) There exists a function $F^\infty : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ such that*

$$V^R(\mathcal{H}_t) = F^\infty(G_t) - G_t. \quad (4)$$

¹²In an extension (Appendix D), c is allowed to also depend on \bar{B}_t , thus capturing the idea that current technologies affect the cost of further radical experimentation.

(ii) The value of radical experimentation is decreasing in the gap G_t . (iii) If $V^R(\mathcal{H}_t)$ is negative in some period t , radical experimentation is abandoned forever after.

Proposition 1 shows that the value of radical experimentation depends on history only through the gap G_t : distinct histories inducing the same gap yield identical incentives for radical experimentation. In particular, the payoff level of the best explored technology does not matter *per se*. Accordingly, the value of radical experimentation at time t will be denoted more simply by $V^R(G_t)$.

The incentives for radical experimentation are driven by the probability of finding a new technology whose payoff exceeds \bar{B}_t . Since the payoff distribution of new technologies is pinned down by the payoff at the frontier, \hat{B}_t , a higher gap reduces the probability of this event. The monotonicity property (ii) implies the stagnation result contained in (iii): By definition, the value of radical experimentation is negative if and only if exploiting the best known technology strictly dominates radical experimentation. Therefore, if the value of radical experimentation becomes negative, it can never be positive again, because \hat{B}_t is frozen, while \bar{B}_t can only increase. The proposition thus has consequences for the sustainability (or lack thereof) of radical experimentation in the long run, which are exploited in later sections.

The next result characterizes the optimal size of radical experimentation, and shows that the *marginal* value and the size of radical experimentation are both decreasing in the highest known payoff, \bar{B}_t .

Let $x^R(\mathcal{H}_t)$ denote the technology chosen when radical experimentation takes place and $y_t^R(\mathcal{H}_t) = x^R(\mathcal{H}_t) - \hat{x}_t$ denote the optimal size of radical experimentation.¹³ Also let $\phi(\cdot)$ denote the density function of the standard normal distribution.

PROPOSITION 2 (RADICAL EXPERIMENTATION) *Suppose that radical experimentation is optimal. Then, (i) the optimal size y_t^R solves*

$$\underbrace{\frac{\delta\sigma}{2\sqrt{y}}\phi\left(\frac{G_t}{\sigma\sqrt{y}}\right)}_{\text{Marginal Benefit}} = \underbrace{c'(y)}_{\text{Marginal Cost}} \quad (5)$$

(ii) y_t^R is strictly increasing in σ and δ . (iii) If $G_t = 0$, then y_t^R is indeed strictly positive. (iv) If $G_t > 0$ and $y_t^R > 0$, then y_t^R is strictly decreasing in G_t . (v) Finally, there exists a cutoff $\tilde{G} > 0$ above which $V^R(G) < 0$.

To get some intuition for Proposition 2, recall that radical experimentations have an expected payoff of \hat{B}_t , regardless of their size. A larger size provides a higher variance, and thus an increased probability of exceeding \bar{B}_t . With a zero gap, the marginal benefit of radical experimentation

¹³There may exist several optima, although such case happens with zero probability. In such knife-edge cases, the comparative statics in the proposition still apply in the sense of the strong set order of lattice theory.

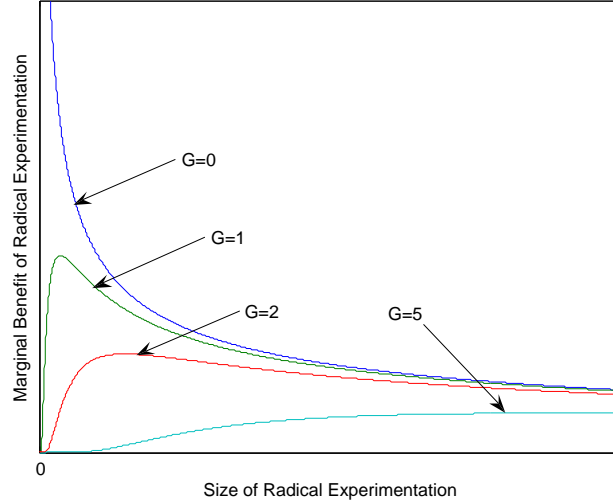


Figure 1: Marginal benefit of radical experimentation for a discount factor $\delta = 0.8$ and volatility $\sigma = 3$.

(left-hand side of (5)) is then arbitrarily large close to the frontier (see Figure 1). As the size of radical experimentation increases, the volatility of the payoff, $\sigma\sqrt{x - \hat{x}_t}$, increases at a decreasing rate, and the marginal benefit of radical experimentation converges to zero. Since the marginal cost is increasing, the optimal size of radical experimentation is therefore positive and well-defined.

The situation is quite different when the gap is positive, because the marginal benefit of radical experimentation now converges to zero for radical experimentations close to the frontier. Close to the frontier, a radical experimentation is accompanied by an inadequately low increase in volatility, and virtually no impact on the expected payoff of the agent when he becomes old. The “outside” option, \bar{B}_t , is thus strictly preferred to small radical experimentations. The marginal benefit of radical experimentation is single peaked: it is initially pushed up by the increase in the probability of discovering a payoff above the current outside option \bar{B}_t , which is given by $1 - \Phi\left(\frac{G_t}{\sigma\sqrt{x - \hat{x}_t}}\right)$, where Φ denotes the distribution function of a standard normal distribution. When the size of radical experimentation reaches $\frac{G_t^2}{\sigma^2}$, the marginal benefit starts to decrease, as the probability of obtaining a payoff greater than \bar{B}_t converges to $\frac{1}{2}$. Figure 1 illustrates the marginal benefit of radical experimentation for different sizes of the gap.

Because volatility is the agent’s only hope of improving the best technology, a higher volatility has a positive effect on the incentives to innovate. Similarly, a higher discount factor increases the incentives to innovate. By contrast, an increase in the gap reduces the marginal benefit of radical experimentation, because it reduces the probability of surpassing the current outside option. To maintain that probability, an agent must increase the size of radical experimentation, so as to

increase payoff volatility.

Finally, if radical experimentation is disappointing enough, so that the gap exceeds some threshold \tilde{G} , then the value of radical experimentation is negative. This, combined with Proposition 1, implies that radical experimentation is abandoned forever.

4 Incremental Experimentation

Analyzing incremental experimentation presents a particular challenge, owing to the correlation across payoff technologies. Experimenting with a new technology reveals information not only about that technology, but also about nearby technologies. Our strategy is to partition the technology space into ‘units’, separated by previously tried technologies. The partition consists of the finitely many bounded intervals whose endpoints have been explored and whose interior points have not, and of the unbounded interval $[\hat{x}_t, \infty)$. A *unit* is defined by any such interval, along with the values of its endpoint payoffs. The distribution of payoffs within a given unit is described by a Brownian bridge: it is the distribution of Brownian motion on some interval with known end values.¹⁴ In particular, it is conditionally independent of the payoff of all observed technology-payoff pairs outside of the unit. Thus, units have the appealing property that experimentation with a technology lying inside a unit changes only the payoff distribution of technologies that belong to that same unit.

The optimization problem of a young agent can be decomposed in two steps: First, choose one of the finitely many units corresponding to history \mathcal{H}_t . Second, choose which technology to pick within that unit. The advantage of this decomposition is that the value of each unit may be characterized according to a simple index, described by Proposition 3. The index determines the *value of experimentation* of the unit, which forms a key block in the analysis of the interplay between radical and incremental experimentation performed in the next section.

For any bounded unit \mathbf{u} , let B^L (B^H) denote the smaller (larger) of the two endpoint payoffs, L the width of the underlying interval, and $D = B^H - B^L \geq 0$ the difference between endpoint payoffs. These variables clearly depend on the unit they are attached to, but the reference to \mathbf{u} is omitted when there is no ambiguity. The unbounded unit at time t is denoted \mathbf{u}^∞ , again omitting the reference to t when there is no ambiguity, which has already been analyzed in the previous section. Figure 2 represents an experimentation history up to $t = 3$.

PROPOSITION 3 *To each bounded unit \mathbf{u} corresponds a value of experimentation, $V(\mathbf{u}, \bar{B}_t)$, with the following properties:*

- (i) *It is optimal for the young agent to choose a technology within the unit (bounded or unbounded) that has the highest value of experimentation.*

¹⁴See Billingsley (1968, p. 64) for an introduction to Brownian bridges.

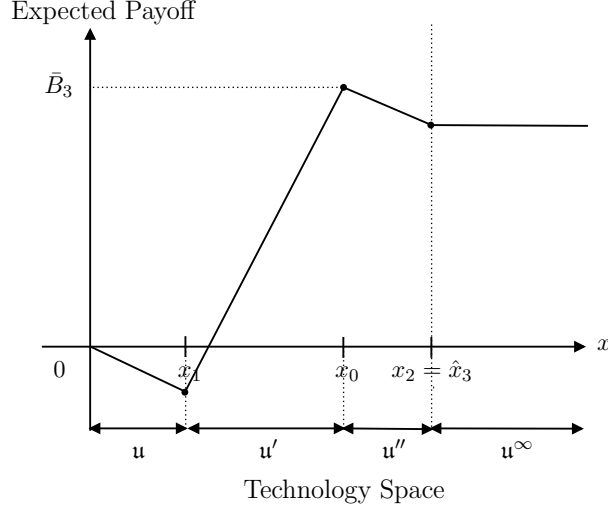


Figure 2: History of units after three periods.

- (ii) There exists a function $F : \mathbb{R}_+^2 \rightarrow [1, \infty)$ such that the value of experimentation of a bounded unit can be written as

$$V(\mathbf{u}, \bar{B}_t) = (B^L - \bar{B}_t) + D F\left(\frac{\sqrt{L}}{D}, \frac{\bar{B}_t - B^L}{D}\right). \quad (6)$$

- (iii) If $B^H < \bar{B}_t$, then $V(\mathbf{u}, \bar{B}_t)$ is increasing in L , B^L and B^H .

Moreover, it is strictly optimal, within a bounded unit, to choose a technology closer to the endpoint with the higher payoff.¹⁵

The value of experimentation of a unit \mathbf{u} is similar to the value of radical experimentation. The only difference is that a young agent's choice in the first period is now restricted to a technology within \mathbf{u} , instead of $[\hat{x}_t, +\infty)$. With this restriction, a young agent's value function may be written as $(1 + \delta) [\bar{B}_t + V(\mathbf{u}, \bar{B}_t)]$.

Some units may have a negative value of experimentation. Choosing from those units is always suboptimal: receiving the best explored payoff \bar{B}_t in both periods dominates such choices. The next proposition shows a stronger result: the value of experimentation of a unit is nonincreasing in \bar{B}_t . Because \bar{B}_t is nondecreasing over time, this implies that once a unit gets a negative value of experimentation, it is abandoned forever.

PROPOSITION 4 (VALUE MONOTONICITY) *For any bounded unit \mathbf{u} , $V(\mathbf{u}, \bar{B}_t)$ is strictly decreasing in \bar{B}_t . If, at any time, $V(\mathbf{u}, \bar{B}_t) < 0$, no technology in \mathbf{u} is ever chosen after time t .*

¹⁵As a consequence, it is optimal to choose the midpoint of a unit whose endpoints have the same payoff.

This monotonicity property is shown in two steps. Consider, first, a unit \mathbf{u} that does *not* contain the best explored technology, and suppose that the payoff of that technology is increased from \bar{B}_t to $\bar{B}'_t > \bar{B}_t$. The increase has no effect on the payoff distribution of technologies inside \mathbf{u} , but it reduces the probability that any technology in \mathbf{u} beats the best explored technology. This, intuitively, reduces the value of experimenting in that unit, implying that $V(\mathbf{u}, \bar{B}'_t) < V(\mathbf{u}, \bar{B}_t)$. Now, consider a unit \mathbf{u} whose endpoints include the best explored technology. For such a unit, a higher value of \bar{B}_t increases, linearly, the expected payoff of all technologies inside the unit. The variance of payoffs within that unit is unaffected (this is a standard property of Brownian bridge), however, so exploitation with the best explored technology is relatively more appealing than before the increase, which again reduces the value of experimentation.¹⁶

A common issue in experimentation and social learning models is whether, and with what probability, agents converge in finite time to some specific action. The next result shows that *exploitation* – choosing a known technology – is strictly suboptimal, which implies that a new unit is created at each period and, hence, that the partitioning in units of the technology space becomes strictly finer over time.

PROPOSITION 5 (EXPLOITATION) *If \mathbf{u} contains the best explored technology, then $V(\mathbf{u}, \bar{B}_t) > 0$.*

This proposition implies that choosing \bar{B}_t for both periods is strictly dominated by choosing another technology included in a unit \mathbf{u} whose endpoint payoffs include \bar{B}_t . Intuitively, a slight departure from the best explored technology reduces the expected payoff of the agent, but it also creates an option value that dominates that reduction, because volatility increments of Brownian motion (of order \sqrt{dt}) dominate expectation increments (of order dt). By the same argument, exploitation would remain suboptimal *even if incremental experimentation were costly*, as long as the cost of incremental experimentation goes smoothly to zero as one gets closer to known technologies. This property is convenient for the analysis: it implies that one does not need to separate pure exploitation from experimentation. In substance, however, one should note that there is very little difference, in this model, between pure exploitation and a very small amount of experimentation: the payoffs and amount of information are very similar in both cases, with very high probability.

To distinguish between incremental and radical experimentations, define the **value of incremental experimentation**, $V^I(\mathcal{H}_t)$, as the maximum value of experimentation over all bounded

¹⁶A starker intuition for this result can be obtained by appealing to the theory of large deviations (see, Dembo and Zeitouni 1998). As \bar{B}_t gets arbitrarily large, the payoff distribution inside the unit \mathbf{u} looks closer to a straight line, joining the low-payoff extremity x_l to the best technology x_r with payoff \bar{B}_t : $(B(x) - B(x_l))/(\bar{B}_t - B(x_l)) \rightarrow_{\bar{B}_t \rightarrow \infty} B^L + (\bar{B}_t - B^L)(x - x_l)/(x_r - x_l)$ a.s., where B^L is the payoff at x_l . As \bar{B}_t gets arbitrarily large, therefore, the probability that any given technology x in \mathbf{u} surpasses \bar{B}_t , converges to zero. Exploitation of x_r remains suboptimal for all values of \bar{B}_t , however, as guaranteed by Proposition 5.

units. Given the concept of unit, the value of radical experimentation can be alternatively defined as the value of experimentation of the unbounded unit: $V^R(\mathcal{H}_t) = V(\mathbf{u}^\infty(\mathcal{H}_t), \bar{B}_t)$.

From Proposition 3, *an agent prefers radical over incremental experimentation if and only if the value of radical experimentation exceeds the value of incremental experimentation.*

5 The Dynamics of Social Experimentation

In this model, social experimentation takes the form of experimentation cycles, characterized in this section, in which radical experimentation is followed by incremental experimentation. Radical experimentation generates a positive externality on all future generations by expanding the frontier of explored technologies and, hence, the set of incremental experimentations.

The set $[0, \hat{x}_t]$ of technologies that are convex combinations of previously explored ones will be referred to as the *active domain*. While incremental experimentation refines knowledge about technologies in the active domain, it is characterized by more predictable outcomes and lacks upside potential. As the learning value of incremental experimentation goes down, radical experimentation becomes attractive again, provided that the value of radical experimentation remains positive. Following a highly successful radical experimentation, the best explored technology lies at the frontier. In that case, further radical experimentation is equally likely to outperform or underperform the payoff of the current frontier, and a wave of radical experimentations can occur. This and other results are formalized in Proposition 6.

PROPOSITION 6 (RADICAL VERSUS INCREMENTAL EXPERIMENTATION) *Consider a history \mathcal{H}_t for which $V^R(G_t) > 0$. Then, experimentation has the following properties:*

1. *The probability that radical experimentation occurs at time t or at some future date is strictly positive, regardless of the value of incremental experimentation.*
2. *Suppose that $G_t = 0$.¹⁷ If radical experimentation takes place at time t , there exist (history-dependent) cutoffs b_t^R and b_t^I such that, letting \hat{B}_{t+1} and \mathbf{u}_t denote the payoff resulting from that radical experimentation and, respectively, the newly created unit (based on the interval between old and new boundaries),*
 - i) radical experimentation is optimal at time $t + 1$ if and only if $\hat{B}_{t+1} > b_t^R$, and*
 - ii) if incremental experimentation is optimal at time $t + 1$, it takes place in the newly created unit \mathbf{u}_t if and only if $\hat{B}_{t+1} > b_t^I$.*

¹⁷In the Appendix, we analyze both the case in which $G_t = 0$ and $G_t > 0$. Here, however, we restrict attention to a zero gap for expositional clarity.

Thus, highly successful radical experimentation begets more radical experimentation, while moderately successful experimentation is followed by incremental experimentation.¹⁸ To appreciate the subtlety of i), one may observe that a high payoff increases *both* the value of incremental experimentation in the new unit and the value of radical experimentation. Why does the latter dominate the former? Intuitively, the conditional payoff distribution on the newly created unit is, for very high payoff realizations, roughly a straight line, with very low variance.¹⁹ This means that technologies in the new unit have a lower expectation than the frontier technology, and a low variance. By contrast, radical experimentation has the same expectation as the value of the frontier technology, and a variance that is independent of that level, which makes it more attractive. Whether incremental experimentation occurs within the newly created unit, however, depends on whether or not the outcome of radical experimentation performed at time t , \hat{B}_{t+1} , exceeds the critical threshold b_t^f . Otherwise, incremental experimentation, when optimal, will take place in the old domain.

The previous results may be summarized as follows: From Proposition 5, we know that exploitation is never optimal. Thus, any wave of incremental experimentations reduces the width of units in the active domain. This, all else equal, reduces the value of experimentation in the active domain, by Proposition 3. The value of incremental experimentation may then decrease to the point of triggering a new round of radical experimentation, which continues until it leads to a disappointing payoff, triggering a new wave of incremental experimentations to explore further the units created within the new frontier, giving rise to alternating cycles between radical and incremental experimentation.

Because the technological domain is unbounded, it is a priori possible that agents indefinitely expand the boundary of the active domain as time goes by. Such qualitative property of the model would correspond to unlimited radical experimentation, and to a never-ending succession of radical and incremental experimentation cycles. By contrast, without radical experimentation, technology and knowledge converge to finite levels. Our next result shows that unlimited radical experimentation never occurs. This result persists even when radical experimentation entails an arbitrarily positive drift, as long as the optimal policy is well-defined.²⁰

THEOREM 1 *Radical experimentation ends in finite time with probability one. After radical experimentation has ended, the value of incremental experimentation converges to zero almost surely.*

Thus, only incremental experimentation persists along the equilibrium path. As the value of

¹⁸These qualitative features of the experimentation process clearly remain valid with a positive gap.

¹⁹This result is well-known in the literature on large deviations, see e.g., Dembo and Zeitouni (1998). See also Footnote 16.

²⁰This result is shown in Section B of the online appendix.

additional experimentation goes to zero, we observe the emergence of a technological standard in the limit. This is because all limiting technologies (when more than one limit technology exists) must be payoff-equivalent almost surely.²¹

A key step for understanding and proving Theorem 1 is to establish comparative statics for the incentives of radical experimentation, which we performed in Section 3. The rest of the argument applies ideas from the experimentation literature to analyze our spatial model of technological space.

Proposition 2 hints at the reason why radical experimentation cannot be sustained in the long run: for high gaps, the optimal size of radical experimentation drops, which reduces volatility and, hence, the value of radical experimentation, just when volatility is most needed to make radical experimentation attractive. That proposition alone, however, does *not* imply that radical experimentation terminates in finite time, because the technologies observed in equilibrium are chosen endogenously. To prove stagnation, a key step is to show that, if the value of radical experimentation were positive at all times, the frontier would keep expanding in steps that are bounded below away from zero.²² Intuitively, the size of radical experimentation is decreasing in the gap and we know from Proposition 2 that a high enough gap ends radical experimentation forever. For radical experimentation to continue, therefore, the size of the radical experimentation must be bounded below. However, radical experimentation will eventually hit a region of the technology space where the frontier payoff is so disappointing that the resulting gap exceeds the critical cutoff \tilde{G} identified in Proposition 2.

Figure 3 shows a simulated path for the technology payoff function B , along with the equilibrium experimentation dynamics for two cost specifications. The figure highlights several interesting features of our model. Experimentation cycles arise endogenously along the equilibrium path. Both the length and the number of cycles is path dependent. In particular, it is not necessarily the case that a lower cost of radical experimentation leads to more radical experimentation. The figure shows that radical experimentation may indeed end sooner when experimentation is cheaper. Cheaper experimentation leads to a larger size of radical experimentation when the gap is zero but this larger size may lead an agent to explore an unattractive part of the technology space which is exactly what happens in the simulation. In general, a wave of radical experimentations can go on for several periods before radical experimentation ends forever. Eventually, the search for a better technology begins to cluster around a technology discovered by a previous generation,

²¹This result bears some resemblance to the *informational cascades* analyzed in the social learning literature, especially according to the definition provided by Lee (1993) for the case of a continuum of actions. As we already mentioned in the introduction, the mechanism is different here, because the amount of information is endogenously acquired by each generation.

²²See Lemma 2.

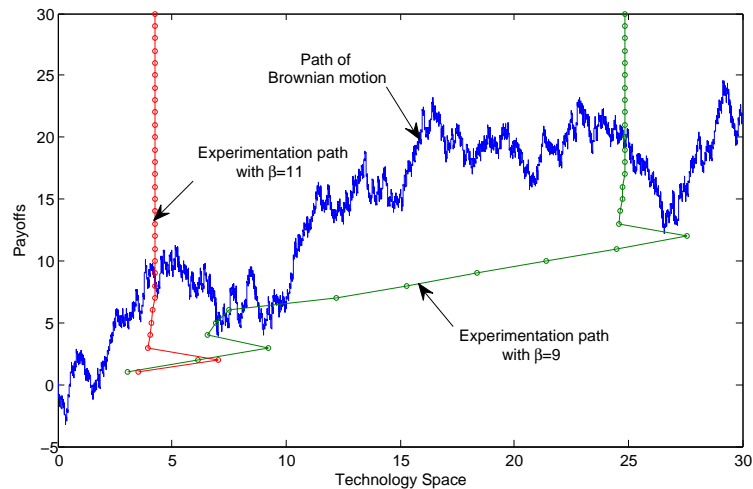


Figure 3: Dynamics of experimentation: $\delta = 1$, $c(y) = \frac{y^2}{2\beta}$. The payoff function is the realized path of a Brownian motion with zero drift and volatility $\sigma = 3$.

yielding the endogenous-information equivalent of an informational cascade. Society converges to a suboptimal technology, which is largely path dependent: a single experiment at, say, $x = 10$ would completely change the dynamics of the search process when radical experimentation is cheaper, shifting experimentation towards a different part of the technological space which would uncover better technologies.

6 Stagnation and Social Inefficiency

Our stagnation result points to a general inefficiency problem with social experimentation. Radical experimentation generates two positive externalities on future generations, by increasing knowledge and by expanding the active domain. If these externalities were taken into account by current generations, wouldn't they systematically push radical experimentation beyond the equilibrium level characterized in earlier sections? We formalize this question by introducing an infinitely-lived social planner who discounts payoffs with a discount factor $\delta_S < 1$.

Comparing the equilibrium dynamics of experimentation with the social optimum raises several difficulties. Firstly, it is well-known that optimal experimentation with an infinitely lived agent and correlated technologies cannot be characterized by the arm-specific index policies that Gittins and Jones (1974) have identified for the standard multi-armed bandit problem with independent arms.²³ This is intuitive: trying a new technology teaches something about surrounding technologies and

²³See Appendix E for an example.

affects the set of units.²⁴ Secondly, any comparison between equilibrium and efficient experimentation has to be of a probabilistic nature. Indeed, it is easy to construct specific Brownian paths for which equilibrium experimentation will last longer than the social optimum, by assigning an implausibly low payoff to the first technology tried by the social planner and, repeatedly high payoffs for the radical experimentations arising in equilibrium. The resulting histories are consistent with possible Brownian paths, and clearly result in equilibrium radical experimentations lasting longer than what is prescribed by the efficient policy. Thirdly, it is not clear how the social planner should discount, if at all, future generations.

We circumvent all those difficulties by considering a limiting result as δ_S goes to 1, i.e., when the social planner becomes infinitely patient. Let $\hat{x}^{FB}(\delta_S)$ denote the frontier at which a social planner with discount factor δ_S stops radical experimentation. Despite our inability to characterize the social optimum notwithstanding, we can prove the following result.

THEOREM 2 *As δ_S goes to 1, $\text{Prob}(\lim_{\delta_S \rightarrow 1} \hat{x}^{FB}(\delta_S) = +\infty) = 1$.*

Theorem 2 implies, as a corollary, that equilibrium experimentation is with high probability inefficiently low compared to the social optimum.

7 Beyond Two-Period Agents

Our earlier focus on two-period agents allowed us to tractably analyze experimentation with interdependent technologies and emphasize how past experimentation affects the option value of new technologies. This section discusses whether the insights developed in the baseline model extend to agents living for more periods. In particular, does stagnation still occur in this case? Longer-lived agents have stronger incentives to experiment, in a sense that we will quantify, and the possible effects of early experimentation on future periods are more complex than with two-period agents. To maintain tractability, we mainly focus on a modification of the baseline in which agents live for three periods and each generation is born in the last (third) period of the previous generation.

To avoid confusion, we will refer to an agent in his first period as *newly born*, while an agent in his second and third period will be referred to as *young* or *old*, respectively, as in the baseline model. Given a discount factor $\delta > 0$, period index $T = 0, 1, 2$ and history \mathcal{H}_t , let $U^I(\mathcal{H}_t; T)$ and $U^R(\mathcal{H}_t; T)$ denote the value functions of an agent with $2 - T$ periods left in the future when his time- t choice is restricted to incremental and radical experimentation, respectively. Recalling the

²⁴Readers may wonder whether units, which are conditionally independent, can be characterized by Gittins-like indices. We have explored this possibility and concluded that this was highly unlikely: an index of the kind that we obtained for two-period lived agents does not seem to exist for infinitely lived agents.

convention that $c(y) = 0$ for any $y \leq 0$, we have

$$U^I(\mathcal{H}_t; T) \equiv \sup_{\{x_{t+s}\}_{s=0}^{2-T}} E_{\mathcal{H}_t} \left[B(x_t) + \sum_{s=1}^{2-T} \delta^s (B(x_{t+s}) - c(x_{t+s} - \hat{x}_{t+s})) \right], \quad (7)$$

s.t. $x_t \in [0, \hat{x}_t], x_{t+s} \in \mathbb{R}_+, s = 1, \dots, 2 - T$

and

$$U^R(\mathcal{H}_t; T) \equiv \sup_{\{x_{t+s}\}_{s=0}^{2-T}} E_{\mathcal{H}_t} \left[B(x_t) - c(x_t - \hat{x}_t) + \sum_{s=1}^{2-T} \delta^s (B(x_{t+s}) - c(x_{t+s} - \hat{x}_{t+s})) \right]. \quad (8)$$

s.t. $x_t \geq \hat{x}_t, x_{t+s} \in \mathbb{R}_+, s = 1, \dots, 2 - T$

The values of incremental and radical experimentation are defined as in the baseline model by comparing the agent's restricted value function to what he would get by exploiting the best known technology at all periods:

$$\mathcal{V}^I(\mathcal{H}_t; T) \equiv U^I(\mathcal{H}_t; T) - \sum_{s=0}^{2-T} \delta^s \bar{B}_t, \quad \text{and} \quad \mathcal{V}^R(\mathcal{H}_t; T) \equiv U^R(\mathcal{H}_t; T) - \sum_{s=0}^{2-T} \delta^s \bar{B}_t. \quad (9)$$

From this simple normalization, it follows that the agent prefers incremental over radical experimentation at history \mathcal{H}_t if and only if $\mathcal{V}^I(\mathcal{H}_t; T) \geq \mathcal{V}^R(\mathcal{H}_t; T)$.

The cases $T = 1$ and $T = 2$ correspond to the young and old agents already analyzed in our baseline model. We now consider the case $T = 0$ of a newly born agent. The value of radical experimentation can be written recursively as²⁵

$$\begin{aligned} \mathcal{V}^R(\mathcal{H}_t; 0) = \sup_{x_t \geq \hat{x}_t} E_{\mathcal{H}_t} [& B(x_t) - c(x_t - \hat{x}_t) - \bar{B}_t + \delta \max\{\mathcal{V}^I(\mathcal{H}_{t+1}; 1), \mathcal{V}^R(\mathcal{H}_{t+1}; 1)\} \\ & + \delta(1 + \delta) (\max\{\bar{B}_t, B(x_t)\} - \bar{B}_t)]. \end{aligned} \quad (10)$$

The main challenge is to characterize the marginal value of radical experimentation for a newly born agent, since it is this value which, compared with the marginal cost, determines the size of radical experimentation. This marginal value turns out to be significantly more complicated with three periods, not least because radical experimentation in the first period can now create new incentives for radical experimentation in the second period of a 3-period agent. This effect was absent with two periods because the agent was already old after performing a radical experiment.

Radical experimentation by a newly born agent affects both the value of radical experimentation and the value of incremental experimentation in the agent's second (young) period. First, it creates a new unit which becomes available for incremental experimentation when the agent is young. Second, a radical experiment also affects the size of the gap in the following period. It is for example

²⁵A similar recursive expression holds for the value of incremental experimentation.

possible that the size of the gap when the agent is young is smaller than when the same agent was newly born due to a successful radical experiment. Thus, the value of radical experimentation for a young agent may be increased, compared to the first period, as an outcome of first period experimentation.

Nonetheless, a newly born agent has a stronger incentive to perform radical experimentation than a young agent, all else equal.²⁶

PROPOSITION 7 *For any history \mathcal{H}_t of tried technologies and corresponding outcomes, an agent's value of radical experimentation is decreasing in his age: $\mathcal{V}^R(\mathcal{H}_t; 0) \geq \mathcal{V}^R(\mathcal{H}_t; 1) \geq \mathcal{V}^R(\mathcal{H}_t; 2) = -G_t$.*

While intuitive, this result is challenging to establish due to the various implications of radical experimentation by a newly-born agent.

The interplay between radical and incremental experimentation is also more complex, and the value of radical experimentation for a newly-born agent is no longer simply determined by the gap $\bar{B}_t - \hat{B}_t$. Nevertheless, an increase in the highest-known payoff, which increases the gap, still reduces the value of radical experimentation, all else equal. Moreover, if the value of radical experimentation becomes negative, then all future generations still discard radical experimentation.

PROPOSITION 8 *(i) The value of radical experimentation for a newly born agent is decreasing in the size of the highest known payoff, \bar{B} , all else equal. (ii) If the value of radical experimentation is negative at some time t , radical experimentation is abandoned by all current and future agents.*

The optimal radical experiment, $x_t^{R,0}(\mathcal{H}_t)$, for a newly born agent now solves the first-order condition

$$\frac{\delta(1+\delta)\sigma}{2\sqrt{x-\hat{x}_t}}\phi\left(\frac{G_t}{\sigma\sqrt{x-\hat{x}_t}}\right) + \delta\frac{dE_{\mathcal{H}_t}[\max\{\mathcal{V}^I(\mathcal{H}_{t+1}; 1), \mathcal{V}^R(\mathcal{H}_{t+1}; 1)\}|x]}{dx} = c'(x - \hat{x}_t), \quad (11)$$

which generalizes equation (5). The first term of (11) is the immediate benefit of a bolder radical experimentation on the probability of exceeding the current best payoff \bar{B}_t , through a higher variance. This effect is stronger than in the baseline model because the benefit of the current radical experiment extends over a longer horizon. The second term of (11) represents the *intertemporal* benefit of bolder radical experimentation, and was absent from the baseline model. It captures the positive effect that a radical experiment today generates on the experimentation opportunities of the agent one period ahead. While the intertemporal marginal benefit boosts incentives to perform radical experimentation for a newly born agent, its effect can be showed to be uniformly bounded above. This observation is used to show the existence of a gap threshold above which radical experimentation becomes suboptimal.

²⁶For the results presented in this section, we assume that the value functions are differentiable.

PROPOSITION 9 (i) If $G_t = 0$, then $y^{R,0}(\mathcal{H}_t) > 0$. (ii) There exists a value of the gap \hat{G} such that $\mathcal{V}^R(\mathcal{H}_t; 0) < 0$ whenever $G_t > \hat{G}$.

Just like radical experimentation, incremental experimentation becomes harder to characterize with longer-lived agents. The two-period equation (6) does not have a simple equivalent with three periods. Some insights developed of the baseline do extend to a longer life-span. For example, exploitation is suboptimal for a newly born agent, just as it was suboptimal for a young agent. Moreover, there still is a positive probability that radical experimentation occurs at or after any history where the value of radical experimentation is positive (Proposition 6.1 in the baseline model).

New issues also arise. For example, successful radical experimentation by a newly born agent replenishes the value of further radical experimentation, but the same agent, when young, may prefer to switch to an incremental experiment in the newly created unit. In addition, the short-run dynamics of experimentation are now affected by the life cycle of new generations: each new generation brings in higher incentives to experiment radically due to an exogenously longer life-span compared to the aging previous generation.

Despite this additional complexity, it is possible to prove that stagnation surely arises in the long run.

THEOREM 3 *Consider the extended model with three-period-lived agents. Radical experimentation ends in finite time with probability one.*

In summary, the trade-off between radical and incremental experimentation is more complex than in the case of two-period agents, but the long run dynamics look the same. While the size of the gap needed to induce a negative option value of radical experimentation may be larger with a longer life-span, past successful incremental experimentation or disastrous radical experimentation will eventually extinguish any incentive to perform radical experimentation due to the path-dependence of the process of experimentation.

Generations overlapping for more than one period

Suppose now that new generations are born when the previous generation is still young. Possible strategic effects are free riding and encouragement effects. Free riding might occur because an agent can avoid radical experimentation in the hope that some other agent will incur the cost of expanding the current frontier. Radical experimentation pursued by one agent may, however, also encourage other agents to pursue more radical experimentation when previous experimentation led to the discovery of an attractive technology. Nevertheless, the long-run dynamics of experimentation would still be affected by large failures at the frontier which should cause the end of radical

experimentation once the gap exceeds a critical level. The short-run dynamics of social experimentation will instead be affected by free-riding and encouragement effects. While clearly beyond the scope of this paper, these effects could in principle be investigated through the way in which they affect the option values of radical and incremental experimentation.²⁷

In our framework, new strategic issues would also arise, beyond free-riding and encouragement effects. Consider the experimentation game between a young agent and a newly born agent. Suppose that the newly born agent experiments radically. From the point of view of the young agent, a smaller radical experiment is less valuable than what it would be in the absence of the other agent's radical experiment. This is due to two factors. First, an additional radical experiment involves a duplication of costs. Second, a larger radical experiment has a larger option value of experimentation. Therefore, the young agent may prefer to experiment incrementally this period and use the information generated by the other agent when old. Similarly, a radical experiment by the young agent may also reduce the option value of radical experimentation for the newly born agent even if such an agent might have performed a larger radical experiment had he been the only active agent. Thus, a multiplicity of equilibria may follow.

8 Discussion

8.1 Experimentation Dynamics: An Example

The history of fuel efficiency illustrates both socially inefficient experimentation and the negative relationship between past successes and incentives for radical experimentation. After the oil crisis of the 1970s, fuel efficiency for passenger vehicles doubled within ten years, but it has stayed roughly at the same level for nearly 20 years.²⁸ Past successful efforts to increase fuel efficiency have channeled further investments toward improvements of the same technologies, instead of pushing the search for new radical technologies. There is, however, a broad agreement that radical experimentation is required to significantly increase fuel efficiency, which is close to its theoretical upper bound under current technologies. While the myopic benefits associated with the exploration of radical technologies may be low compared to the presumably high development costs, the unquestionable

²⁷With infinitely-lived agents, some of these effects have been studied by Bolton and Harris (1999), Keller, Rady, and Cripps (2005), and Keller and Rady (2010). Akcigit and Liu (2013) identifies another interesting aspect of strategic experimentation, where agents hide their research failures, resulting in inefficient replication of unsuccessful research. The limitation of those models is that they focus exclusively on two-arm settings with independent arms. While we restrict attention to forward-looking but non-strategic behavior, we complement the literature by enriching the action space to encompass a continuum of technologies with interdependent and unbounded payoffs.

²⁸See the report "History of Fuel Economy: One Decade of Innovation, Two Decades of Inaction" by the PEW Environment Group, available online at <http://www.pewenvironment.org>.

large social value that such technologies could generate suggests a fitting illustration of inefficient stagnation.

8.2 Breaking Stagnation

The mechanism behind our stagnation result is simple: as radical experiments with uncertain and unbounded payoffs expand the technological frontier, the frontier must at some point get so unattractive relative to the best known technology that no agent wishes to experiment beyond it. This evolution captures Gordon (2012)'s view that radical innovations eventually get stuck as future generations become unwilling to make the necessary investments to further push the technological frontier. It is the willingness of each generation to undertake bold experimentation that fades away over time, not the possibility to innovate radically.

There are two approaches to break the stagnation result: by increasing the new generation's value of radical experimentation, or by decreasing the cost of radical experimentation. Consider, first, the cost reductions. A natural extension is to allow the cost of radical experimentation to depend on known technologies: incremental innovations are helpful to produce more radical ones. In Appendix D, we analyze a setting in which the cost of radical experimentation is decreasing in the value of the best known technology. As long as this cost remains strictly positive and convex, however, we show that stagnation still arises with probability 1. The result persists if we add a positive fixed cost for incremental experimentation, which could a priori help by making radical experimentation more attractive relatively to incremental experimentation.

A more radical extension of the model would feature population growth as in Kortum (1998) or capital accumulation, as in standard growth models (Acemoglu 2008). The cost of any given radical experimentation would then decrease in per capita terms or become small relative to the capital available. These features would make it more affordable for later generations to increase the size of radical experimentation, perhaps to the point at which an unattractive technology at the frontier becomes irrelevant relative to much bolder innovations. This approach would likely suffice to break stagnation as long as the population or capital available grows in each period regardless of the technological choices made by current generations.

Turning to the value of experimentation, we analyze in the online appendix (Section B), a variation of the model in which the Brownian path has an arbitrary positive drift.²⁹ Intuitively, this drift captures the idea that even when the frontier technology is disappointing, the outlook of moving beyond the frontier gets increasingly better with the size of this move. However, we show

²⁹We require that the marginal cost of radical experimentation exceeds this drift as the size the of experimentation becomes arbitrarily large. Without this assumption, the optimal size of radical experimentation may be infinite and the problem is not well defined.

that this channel does not work: stagnation still arises with probability one.

Our model assumes that radical experimentations are building on the current frontier. In reality, there are often multiple directions to follow for radical experimentation, and if the frontier along one direction is not promising, one can always experiment along another direction. With finitely many directions, however, this extension of the model would not help: one would still reach with probability 1 a point at which *all* directions look unappealing.³⁰

Another approach is to consider a fully multidimensional technological space, instead of just multiple one-dimensional technological spaces. A natural modeling choice here would be to use a Brownian sheet, which is a multidimensional extension of Brownian motion (see, Revuz and Yor 1999).³¹ In such a model, each previously tried technology would be informative about the payoff distribution of the new technology. This feature is perhaps realistic, since even remotely connected technologies may be helpful to each other. But this would also significantly complicate the analysis by breaking the Markov property which was instrumental in analyzing the “units” developed in this paper.

In modeling a single decision at each period, we followed the multi-armed bandit and social learning literatures where agents must choose a single bandit or action at each time. A slight modification of this paradigm (but not anodyne for our stagnation result, as discussed above) is to follow Jovanovic and Rob (1990) by allowing the agent to peek at new technologies before deciding which one to choose. This possibility may be unrealistic when projects have large fixed costs. For example, the only way to really learn the value of a large particle collider is to build one. Alternatively, agents may be allowed to split their resources between multiple technologies. For example, the current generation could exploit a known technology to guarantee enough immediate consumption but also invest some of its resources in a more radical experiment. If utility from consumption is concave and the generation is sufficiently long-lived, patient, or altruistic towards future generations, such a model would deliver more radical experimentation. However, this alone would likely not suffice to overturn our stagnation result: agents with a fixed lifespan would always reach a period at which the technology at the frontier is so poor as to deter them from making the radical investment.

In our model, generations do not directly interact with one another. With longer lived generations and transfers across them, patents would present another potential channel to foster radical experimentation. If, in particular, older generations are allowed to patent the technologies devel-

³⁰Jovanovic and Rob (1990) consider a setting with infinitely many such dimensions. Even in their setting, however, stagnation would obtain if agents were not allowed to “peek” at the values of new technologies before choosing their technology.

³¹Brownian sheets have multiple indices. They are not equivalent to n dimensional Brownian motions, and indeed are much more complex mathematical objects.

oped when they were young, this may push young generations to make their own experiment, either to avoid paying licensing fees or in the hope of receiving licensing fees in later periods. These mechanisms are discussed in Section F of the appendix, but they have well-known limits. In particular, the perspective of future experimentation can also reduce today's incentives to experiment, as new technologies quickly become obsolete. The process of creative destruction can be detrimental to radical experimentation through direct and equilibrium effects, as shown by Aghion and Howitt (1992), and also lead to stagnation.

Finally, if each generation consists of multiple agents, competition may also create incentives for experimentation. Our model has assumed that there was a single firm experimenting in each period, and that the payoff of any given technology was exogenous. In practice, however, various firms may compete in the use of any given technology, which may crowd out profit and increases the costs of the relevant resources. This may spur some firms to choose different technologies, assuming that technologies are not perfect substitutes, in order to achieve horizontal differentiation.

As this section illustrates, many interesting questions remain open for successfully incorporating the experimentation paradigm into the innovation and growth literatures. Our paper provides one step in this direction and a framework on which to base further research on the topic.

Appendices

A Proofs of propositions from Sections 3 and 4

Proof of Proposition 1

Given an arbitrary history \mathcal{H} ,³² the expected utility from a radical experimentation $x > \hat{x}$ is

$$\begin{aligned} E[B(x) - c(x - \hat{x}) + \delta \max\{B(x), \bar{B}\}] &= (1 + \delta)\hat{B} - c(x - \hat{x}) \\ &\quad + \delta E \max\{B(x) - \hat{B}, \bar{B} - \hat{B}\} \\ &= (1 + \delta)\hat{B} - c(x - \hat{x}) + \delta E \max\{k(x), G\}, \end{aligned}$$

where $k(x) = B(x) - \hat{B}$. Recall that $k(x) \sim \mathcal{N}(0, \sigma^2(x - \hat{x}))$. Therefore,

$$E \max\{k(x), G\} = \sigma\sqrt{x - \hat{x}}\phi\left(\frac{G}{\sigma\sqrt{x - \hat{x}}}\right) + G\Phi\left(\frac{G}{\sigma\sqrt{x - \hat{x}}}\right). \quad (12)$$

Taking the supremum over $x \geq \hat{x}$, we obtain

$$\begin{aligned} &\sup_{x \geq \hat{x}} \left\{ (1 + \delta)\hat{B} - c(x - \hat{x}) + \delta\sigma\sqrt{x - \hat{x}}\phi\left(\frac{G}{\sigma\sqrt{x - \hat{x}}}\right) + \delta G\Phi\left(\frac{G}{\sigma\sqrt{x - \hat{x}}}\right) \right\} \\ &= (1 + \delta) \left[\hat{B} + \sup_{y \geq 0} \frac{1}{1 + \delta} \left\{ -c(y) + \delta\sigma\sqrt{y}\phi\left(\frac{G}{\sigma\sqrt{y}}\right) + \delta G\Phi\left(\frac{G}{\sigma\sqrt{y}}\right) \right\} \right], \end{aligned}$$

where $y \equiv x - \hat{x}$. Defining

$$F^\infty(G) = \frac{1}{1 + \delta} \sup_{y \geq 0} \left\{ -c(y) + \delta\sigma\sqrt{y}\phi\left(\frac{G}{\sigma\sqrt{y}}\right) + \delta G\Phi\left(\frac{G}{\sigma\sqrt{y}}\right) \right\}, \quad (13)$$

³²We drop the time subscripts from now on to simplify the exposition.

and $V^R(\mathcal{H}) = \left[\hat{B} + F^\infty(G) \right] - \bar{B} = F^\infty(G) - G$ yields the formula claimed by Proposition 1.

Next, we need to evaluate the effect of an increase in the size of the gap, G , on the value of radical experimentation $V^R(G) = F^\infty(G) - G$.³³ For this and other comparative statics results, we will apply differential techniques to some value functions. These techniques can be applied because the relevant functions are always left and right differentiable, a result which follows from Corollary 4 of Milgrom and Segal (2002).³⁴ All the comparative statics arguments to follow can be applied to the right derivative (or, in case of a decrease of the parameter, to the left derivative). For expositional simplicity, we drop reference to the side of the derivative.

If there exists an interior solution $y^* > 0$ to (13) when $G > 0$, then $\frac{\partial V^R(G)}{\partial G} = \frac{dF^\infty(G)}{dG} - 1 = \frac{\delta}{1+\delta} \Phi \left(\frac{G}{\sigma\sqrt{y^*}} \right) - 1 < 0$. If there is no interior solution, then $F^\infty(G) = \frac{\delta}{1+\delta}G$ which gives $\frac{\partial V^R(G)}{\partial G} = \frac{\delta}{1+\delta} - 1 < 0$. When $G = 0$, $V^R(G) = F^\infty(0)$ and $\frac{\partial V^R(G)}{\partial G} = 0$.

Finally, if $V^R(G_t) < 0$, the previous comparative statics, monotonicity of the sequence $\{\bar{B}_t\}$, and the fact that \hat{B}_t is unaffected by incremental experimentation, imply that radical experimentation will never be undertaken for any $t' > t$, as it is always dominated by exploitation of the best explored technology. ■

Proof of Proposition 2

If $G = 0$, the expected utility from radical experimentation of size $y > 0$ is equal to

$$E \left[B(\hat{x} + y) - c(y) + \delta \max\{B(\hat{x} + y), \bar{B}\} \right] = (1 + \delta)\hat{B} - c(y) + \delta\sigma\sqrt{y}\phi(0),$$

from (13). The first-order condition yields Equation (5) with $G = 0$. The right-hand side of (5) is increasing in y , while the left-hand side is strictly decreasing. The left-hand side is also unbounded around 0, and converges to 0 as $y \rightarrow +\infty$. Therefore, there always exists a solution to Equation (5) when $G = 0$, and it is unique. The second-order condition

$$-\frac{\delta\sigma\phi(0)}{4y^{3/2}} - c''(y) < 0,$$

is satisfied, guaranteeing that the first-order condition characterizes maxima.

If $G > 0$, the expected utility from radical experimentation of size $y > 0$ is

$$(1 + \delta)\hat{B} - c(y) + \delta\sigma\sqrt{y}\phi \left(\frac{G}{\sigma\sqrt{y}} \right) + \delta G \Phi \left(\frac{G}{\sigma\sqrt{y}} \right).$$

The first-order condition is again given by (5). Differentiating (5) with respect to G , we obtain $-\frac{\delta G}{2\sigma y^{3/2}} \phi \left(\frac{G}{\sigma\sqrt{y}} \right)$, which is strictly negative. Since the objective function in (13) is submodular in (y, G) , the Strict Monotonicity Theorem 1 of Edlin and Shannon (1998) implies that the optimal size of radical experimentation is decreasing in G . Similarly, the objective function is supermodular in (y, δ) and (y, σ) .

We now show the existence of a cutoff \tilde{G} above which radical experimentation has a negative value. Let $A(y, G) \equiv \frac{\delta\sigma}{2\sqrt{y}} \phi \left(\frac{G}{\sigma\sqrt{y}} \right)$ denote the marginal benefit of radical experimentation, given a gap G and a size y of radical experimentation.³⁵

LEMMA 1 *For any $G > 0$,*

1. $\lim_{y \rightarrow +\infty} A(y, G) = \lim_{G \rightarrow +\infty} A(y, G) = \lim_{y \rightarrow 0} A(y, G) = \lim_{y \rightarrow 0} \frac{\partial A(y, G)}{\partial y} = 0$;

³³As explained in the main text, we adopt the convention of representing the value of radical experimentation directly through its dependence on the gap.

³⁴The value function may fail to be differentiable at parameter values for which there exist multiple maximizers. At such values, however, the value function is still left and right differentiable, with the left (right) derivative being evaluated at the maximizer that minimizes (maximizes) the derivative of the objective with respect to the parameter. See Milgrom and Segal (2002).

³⁵ $A(y, G)$ corresponds to the left-hand side of equation (5).

2. $\frac{\partial A(y, G)}{\partial G} < 0$.

3. $A(y, G)$ is strictly quasi-concave in y , and maximized at $\frac{G^2}{\sigma^2}$.

Proof. The first two limits in 1) and the sign of the derivative in 2) directly follow from the properties of $\phi(\cdot)$. The limit $\lim_{y \rightarrow 0} A(y, G)$ is computed using the fact that $\lim_{y \rightarrow 0} \frac{1}{\sqrt{y}} e^{-\frac{1}{y}} = \lim_{z \rightarrow \infty} \frac{z^{1/2}}{e^z} = 0$. One shows similarly that $\lim_{y \rightarrow 0} \frac{\partial A(y, G)}{\partial y} = 0$. Strict quasiconcavity of A in y comes from the fact that

$$\frac{\partial A(y, G)}{\partial y} = \frac{\delta \sigma}{4y^{3/2}} \left[\frac{G^2}{\sigma^2 y} - 1 \right] \phi \left(\frac{G}{\sigma \sqrt{y}} \right),$$

which is nonnegative below $\frac{G^2}{\sigma^2}$ and negative above. This also shows that $\frac{G^2}{\sigma^2}$ maximizes $A(\cdot, G)$. \blacksquare

To establish the existence of the threshold $\tilde{G} > 0$, we start by showing that $y^* = 0$ is the unique maximizer of (13), whenever G is high enough. If $c'(0) > 0$, Lemma 1 implies that $\lim_{y \rightarrow 0} A(y, G) = 0 < c'(0)$ for any $G > 0$. Also by Lemma 1, $\frac{\partial A(y, G)}{\partial G} < 0$, $\lim_{G \rightarrow +\infty} A(y, G) = 0$, and $A(y, G) \leq A\left(\frac{G^2}{\sigma^2}, G\right) = \frac{\delta \sigma^2}{2G} \phi(1)$. This, combined with the properties of c , implies the existence of a large enough threshold \tilde{G} such that $c'(y) > A(y, G)$ for all $G > \tilde{G}$ and $y > 0$. A similar argument applies if $c'(0) = 0$ and $c''(0) > 0$, because $\lim_{y \rightarrow 0} \frac{\partial A(y, G)}{\partial y} = 0$, from Lemma 1. Substituting $y^* = 0$ into (13) yields $F^\infty(G) = \frac{\delta G}{1+\delta}$. Therefore, the value of radical experimentation is equal to $V^R(G) = -\frac{G}{1+\delta} < 0$, for $G > 0$. \blacksquare

Proof of Proposition 3

Given that the distribution of payoffs within each unit is conditionally independent from other units, fix an arbitrary bounded unit u . Suppose, without loss of generality, that $B^H = B(x_r) > B(x_l) = B^L$ (the case of an equality is obtained by taking the limit as, say, B^L is increased to B^H). Then, for any $x \in [x_l, x_r]$,

$$B(x) \sim \mathcal{N} \left(B^L + \frac{B^H - B^L}{x_r - x_l} (x - x_l), \frac{(x - x_l)(x_r - x)}{x_r - x_l} \sigma^2 \right),$$

where $\mathcal{N}(\kappa, \sigma^2)$ denotes the distribution of a Gaussian variable with mean κ and variance σ^2 . Letting $a(x) = \frac{x - x_l}{x_r - x_l}$,

$$B(x) - B^L \sim \mathcal{N} (D a(x), a(x)(1 - a(x))L\sigma^2).$$

We also define $k(x) = B(x) - B^L$ and $\bar{B}' = \bar{B} - B^L$ to obtain an explicit formula for the expected payoff:

$$\begin{aligned} E \max\{k(x), \bar{B}'\} &= \bar{B}' \Phi \left(\frac{\bar{B}' - Da(x)}{\sigma \sqrt{a(x)(1 - a(x))L}} \right) + Da(x) \left[1 - \Phi \left(\frac{\bar{B}' - Da(x)}{\sigma \sqrt{a(x)(1 - a(x))L}} \right) \right] \\ &+ \sigma \sqrt{a(x)(1 - a(x))L} \phi \left(\frac{\bar{B}' - Da(x)}{\sigma \sqrt{a(x)(1 - a(x))L}} \right), \end{aligned}$$

where Φ and ϕ are the CDF and pdf of the standard normal distribution. This leads to

$$\begin{aligned} E_u [B(x) + \delta \max\{B(x), \bar{B}\}] &= (1 + \delta)B^L + Da(x) + \delta E \max\{k(x), \bar{B}'\} \\ &= (1 + \delta)B^L + D \left\{ a(x) \left(1 + \delta - \delta \Phi \left(\frac{\frac{\bar{B}'}{D} - a(x)}{\sigma \frac{\sqrt{L}}{D} \sqrt{a(x)(1 - a(x))}} \right) \right) \right. \\ &+ \delta \sigma \sqrt{a(x)(1 - a(x))} \frac{\sqrt{L}}{D} \phi \left(\frac{\frac{\bar{B}'}{D} - a(x)}{\sigma \frac{\sqrt{L}}{D} \sqrt{a(x)(1 - a(x))}} \right) \\ &\left. + \delta \frac{\bar{B}'}{D} \Phi \left(\frac{\frac{\bar{B}'}{D} - a(x)}{\sigma \frac{\sqrt{L}}{D} \sqrt{a(x)(1 - a(x))}} \right) \right\}. \end{aligned} \tag{14}$$

Define $q_1 \equiv \frac{\sqrt{L}}{D}$, $q_2 \equiv \frac{\bar{B}'}{D}$, and

$$\begin{aligned} \bar{F}(a, q_1, q_2) = & \frac{1}{1+\delta} \left\{ a \left(1 + \delta - \delta \Phi \left(\frac{q_2 - a}{\sigma q_1 \sqrt{a(1-a)}} \right) \right) \right. \\ & \left. + \delta q_2 \Phi \left(\frac{q_2 - a}{\sigma q_1 \sqrt{a(1-a)}} \right) + \delta \sigma \sqrt{a(1-a)} q_1 \phi \left(\frac{q_2 - a}{\sigma q_1 \sqrt{a(1-a)}} \right) \right\}. \end{aligned} \quad (15)$$

As the common argument of Φ and ϕ depends only on $a(x)$, \sqrt{L}/D and \bar{B}'/D , we can write

$$E_u [B(x) + \delta \max\{B(x), \bar{B}\}] = (1 + \delta) \left\{ B^L + D \bar{F} \left(a(x), \frac{\sqrt{L}}{D}, \frac{\bar{B}'}{D} \right) \right\}. \quad (16)$$

Maximizing the expected payoff over $x \in [x_l, x_r]$ yields

$$U(\mathbf{u}; \mathcal{H}_t) \equiv \max_{x \in [x_l, x_r]} E_u [B(x) + \delta \max\{B(x), \bar{B}\}] = (1 + \delta) \left[B^L + D F \left(\frac{\sqrt{L}}{D}, \frac{\bar{B}'}{D} \right) \right], \quad (17)$$

where

$$F(q_1, q_2) = \max_{a \in [0,1]} \bar{F}(a, q_1, q_2). \quad (18)$$

Letting $V(\mathbf{u}, \bar{B}) = \left[B^L + D F \left(\frac{\sqrt{L}}{D}, \frac{\bar{B}-B^L}{D} \right) \right] - \bar{B}$ yields the characterization of the value of experimentation for a bounded unit.

The value of experimentation of each unit, whether bounded or unbounded, corresponds to a normalization of the agent's value function when the domain of choice is restricted to that particular unit. Therefore, the agent optimally chooses a technology within the unit with the highest value of experimentation.

Next, notice that

$$\begin{aligned} \frac{\partial \bar{F}(a, q_1, q_2)}{\partial a} &= \frac{1}{1+\delta} \left\{ 1 + \delta - \delta \Phi \left(\frac{q_2 - a}{\sigma q_1 \sqrt{a(1-a)}} \right) + \frac{\delta \sigma q_1}{2} \frac{1 - 2a}{\sqrt{a(1-a)}} \phi \left(\frac{q_2 - a}{\sigma q_1 \sqrt{a(1-a)}} \right) \right\} \\ &\geq \frac{1}{1+\delta} > 0, \end{aligned}$$

for any $a \leq \frac{1}{2}$, because the sum of the first three terms in curly brackets is weakly greater than 1 by the property of Φ and the last term is positive if, and only if, $a \leq \frac{1}{2}$. Thus, any $a^* \in \arg \max_{a \in [0,1]} \bar{F}(a, q_1, q_2)$ must be greater than $\frac{1}{2}$. As $a(x) = \frac{x-x_l}{x_r-x_l}$, the corresponding optimal technology from a unit with interval $[x_l, x_r]$, $x^*(\mathbf{u}, \bar{B})$, must lie in $(\frac{x_r+x_l}{2}, x_r]$ whenever $B(x_l) < B(x_r)$.

We suppose once again that the value function (18) is differentiable. The Envelope Theorem implies, for any bounded unit \mathbf{u} , that³⁶

$$\frac{\partial V(\mathbf{u}, \bar{B})}{\partial L} = \frac{\delta}{1+\delta} \frac{\sigma}{2} \sqrt{\frac{a^*(1-a^*)}{L}} \phi \left(\frac{\bar{B}' - D a^*}{\sigma \sqrt{a^*(1-a^*)L}} \right) > 0, \quad (19)$$

where $a^* \in \arg \max_{a \in [0,1]} \bar{F}(a, q_1, q_2)$. Similarly, for $B^H \neq \bar{B}$, $\frac{\partial V(\mathbf{u}, \bar{B})}{\partial B^H} > 0$ and $\frac{\partial V(\mathbf{u}, \bar{B})}{\partial B^L} > 0$, as can be seen from equations (23) and (24) in the proof of Proposition 6. ■

Proof of Proposition 4

For any bounded unit \mathbf{u} with $B^H < \bar{B}$ we have³⁷

$$\frac{dF(q_1, q_2)}{d\bar{B}} = \frac{1}{D} \frac{\partial F(q_1, q_2)}{\partial q_2} = \frac{\delta}{(1+\delta)D} \Phi \left(\frac{q_2 - a^*}{\sigma q_1 \sqrt{a^*(1-a^*)}} \right), \quad (20)$$

³⁶ An argument similar to the proof of Proposition 5, which holds independently of the present comparative statics, shows that any optimum x^* is always in the interior of a unit.

³⁷ Again, we omit dependence on side-derivatives. See Footnote 34 and the discussion surrounding it.

where, as before, $q_1 \equiv \frac{\sqrt{L}}{D}$, $q_2 \equiv \frac{\bar{B}-B^L}{D}$, and $a^* \in \arg \max_{a \in [0,1]} \bar{F}(a, q_1, q_2)$. This implies that

$$\frac{\partial V(\mathbf{u}, \bar{B})}{\partial \bar{B}} = \frac{\delta}{1+\delta} \Phi \left(\frac{\bar{B} - B^L - Da^*}{\sigma \sqrt{a^*(1-a^*)L}} \right) - 1 < 0.$$

If, instead, \mathbf{u} contains the best explored technology, then $V(\mathbf{u}, \bar{B}) = B^L - \bar{B} + (\bar{B} - B^L)F\left(\frac{\sqrt{L}}{\bar{B}-B^L}, 1\right)$. This yields $\frac{\partial V(\mathbf{u}, \bar{B})}{\partial \bar{B}} = F\left(\frac{\sqrt{L}}{\bar{B}-B^L}, 1\right) - \frac{\sqrt{L}}{\bar{B}-B^L} \frac{\partial F}{\partial q_1}\left(\frac{\sqrt{L}}{\bar{B}-B^L}, 1\right) - 1$, which is strictly negative from (25).

Exploitation of the best explored technology is always feasible for a young agent, and yields a payoff of $(1+\delta)\bar{B}_t$. Consider a unit \mathbf{u} such that $V(\mathbf{u}, \bar{B}_t) < 0$. Then,

$$V(\mathbf{u}, \bar{B}_t) < 0 \iff \frac{U(\mathbf{u}; \mathcal{H}_t)}{1+\delta} < \bar{B}_t \iff U(\mathbf{u}; \mathcal{H}_t) < (1+\delta)\bar{B}_t,$$

where the second equivalence follows from (17). Therefore, choosing the best explored technology dominates choosing any technology in \mathbf{u} . Since \bar{B}_t is nondecreasing in t , once the value of experimentation of a unit is negative, it remains negative. \blacksquare

Proof of Proposition 5

Suppose that $B^H = \bar{B}$. Differentiating $\bar{F}(a, q_1, 1)$ with respect to a yields

$$(1+\delta) \frac{\partial \bar{F}(a, q_1, 1)}{\partial a} = \left[1 + \delta - \delta \Phi \left(\frac{1}{\sigma q_1} \sqrt{\frac{1-a}{a}} \right) \right] + \frac{\delta \sigma q_1}{2} \frac{1-2a}{\sqrt{a(1-a)}} \phi \left(\frac{1}{\sigma q_1} \sqrt{\frac{1-a}{a}} \right),$$

which tends to $-\infty$ as a goes to 1. Also, $\bar{F}(0, q_1, 1) \equiv \lim_{a \rightarrow 0} \bar{F}(a, q_1, 1) = \frac{\delta}{1+\delta} < 1 = \lim_{a \rightarrow 1} \bar{F}(a, q_1, 1) \equiv \bar{F}(1, q_1, 1)$. We have thus shown that any solution to (18) lies in $(0,1)$ whenever $q_1 > 0$ and $q_2 = 1$. Consider any history \mathcal{H}_t and unit \mathbf{u} with \bar{B} as one of its endpoint payoffs. Then, $D = \bar{B} - B^L$ and

$$\begin{aligned} V(\mathbf{u}, \bar{B}) &= B^L - \bar{B} + (\bar{B} - B^L)F\left(\frac{\sqrt{L}}{\bar{B}-B^L}, 1\right) > B^L - \bar{B} + (\bar{B} - B^L)\bar{F}\left(1, \frac{\sqrt{L}}{\bar{B}-B^L}, 1\right) \\ &= B^L - \bar{B} + (\bar{B} - B^L) = 0, \end{aligned}$$

which proves that there is always at least one bounded unit with strictly positive value of experimentation. Thus, exploitation is strictly suboptimal. \blacksquare

Proof of Proposition 6

If $V^R(G_t) > V^I(\mathcal{H}_t)$, then radical experimentation occurs with probability 1 at time t . Thus, consider a history \mathcal{H}_t such that $V^I(\mathcal{H}_t) > V^R(G_t) > 0$. By continuity, there exists $\varepsilon > 0$ such that the value of radical experimentation remains positive if the current gap is increased to $G_t + \varepsilon$. Let $\mathcal{B}(\varepsilon)$ denote the set of paths of B on $[0, \hat{x}_t]$ which are compatible with \mathcal{H}_t and are bounded above by $\bar{B}_t + \varepsilon$. $\mathcal{B}(\varepsilon)$ occurs with positive probability. By construction, the sequence $\{\bar{B}_{t'}\}_{t' > t}$ stays below $\bar{B}_t + \varepsilon$ for any path in $\mathcal{B}(\varepsilon)$, which implies that the value of incremental experimentation converges to zero over time, by Theorem 1. Thus, the value of incremental experimentation must fall below the value of radical experimentation, which is uniformly bounded away from zero for any path in $\mathcal{B}(\varepsilon)$.

Next, suppose that radical experimentation occurs at an arbitrary history \mathcal{H}_t . Let \hat{B}_{t+1} denote the corresponding payoff of radical experimentation, and let \mathbf{u}_t denote the newly created bounded unit. For the rest of this proof, it turns out to be more convenient to work with the auxiliary indexes $Q(\mathbf{u}, \bar{B}) = V(\mathbf{u}, \bar{B}) + \bar{B} = B^L + D F(q_1, q_2)$ and $Q(\mathbf{u}^\infty, \bar{B}) = F^\infty(G) + \hat{B}$ which are a simple normalization of the value function of a young agent when his first period choice is restricted to a unit \mathbf{u} . Thus, the young agent experiments with a technology within the unit with the highest (auxiliary) index $Q(\mathbf{u}, \bar{B})$. In particular, we can define the indexes for incremental and radical experimentation by

$$Q^I(\mathcal{H}_t) = \max_{\mathbf{u} \in \mathcal{P}(\mathcal{H}_t)} Q(\mathbf{u}, \bar{B}) \quad \text{and} \quad Q^R(\mathcal{H}_t) = Q(\mathbf{u}^\infty, \bar{B}), \quad (21)$$

for any history \mathcal{H}_t , where $\mathcal{P}(\mathcal{H}_t)$ is the collection of bounded units induced by a history \mathcal{H}_t . After radical experimentation has taken place, the new index for incremental experimentation at the start of time $t + 1$ is

$$Q^I(\mathcal{H}_{t+1}) = \max \left\{ \underbrace{\max_{\mathbf{u} \in \mathcal{P}(\mathcal{H}_t)} Q(\mathbf{u}, \max\{\bar{B}_t, \hat{B}_{t+1}\}), Q(\mathbf{u}_t, \max\{\bar{B}_t, \hat{B}_{t+1}\})}_{\bar{Q}(\max\{\bar{B}_t, \hat{B}_{t+1}\})} \right\}.$$

We first consider how $\bar{Q}(\max\{\bar{B}_t, \hat{B}_{t+1}\})$ varies with \hat{B}_{t+1} . If $\hat{B}_{t+1} \leq \bar{B}_t$, $\bar{Q}(\cdot)$ is unaffected by \hat{B}_{t+1} . If $\hat{B}_{t+1} > \bar{B}_t$, consider any unit $\mathbf{u} \in \mathcal{P}(\mathcal{H}_t)$. Then,³⁸

$$\frac{\partial Q(\mathbf{u}, \hat{B}_{t+1})}{\partial \hat{B}_{t+1}} = \frac{\partial}{\partial \hat{B}_{t+1}} \left[B^L + DF \left(\frac{\sqrt{L}}{D}, \frac{\hat{B}_{t+1} - B^L}{D} \right) \right] = \frac{\partial F}{\partial q_2} \left(\frac{\sqrt{L}}{D}, \frac{\hat{B}_{t+1} - B^L}{D} \right) \in \left(0, \frac{\delta}{1 + \delta} \right), \quad (22)$$

by (20). Next, we consider the index of the new unit \mathbf{u}_t . To this end, recall that \hat{B}_t is the payoff associated with the left endpoint of the newly created unit \mathbf{u}_t . If $\hat{B}_t < \hat{B}_{t+1} \leq \bar{B}_t$, then

$$\begin{aligned} \frac{\partial Q(\mathbf{u}_t, \bar{B}_t)}{\partial \hat{B}_{t+1}} &= \frac{\partial}{\partial \hat{B}_{t+1}} \left[\hat{B}_t + (\hat{B}_{t+1} - \hat{B}_t) F \left(\frac{\sqrt{L}}{\hat{B}_{t+1} - \hat{B}_t}, \frac{\bar{B}_t - \hat{B}_t}{\hat{B}_{t+1} - \hat{B}_t} \right) \right] \\ &= \frac{a^*}{1 + \delta} \left[1 + \delta - \delta \Phi \left(\frac{q_2 - a^*}{\sigma q_1 \sqrt{a^*(1 - a^*)}} \right) \right] > 0, \end{aligned} \quad (23)$$

where $a^* \in \arg \max_{a \in [0, 1]} \bar{F}(a, q_1, q_2)$. If, instead, $\hat{B}_{t+1} < \min\{\bar{B}_t, \hat{B}_t\}$, then

$$\begin{aligned} \frac{\partial Q(\mathbf{u}_t, \bar{B}_t)}{\partial \hat{B}_{t+1}} &= \frac{\partial}{\partial \hat{B}_{t+1}} \left[\hat{B}_{t+1} + (\hat{B}_t - \hat{B}_{t+1}) F \left(\frac{\sqrt{L}}{\hat{B}_t - \hat{B}_{t+1}}, \frac{\bar{B}_t - \hat{B}_{t+1}}{\hat{B}_t - \hat{B}_{t+1}} \right) \right] \\ &= (1 - a^*) \left[1 - \frac{\delta}{1 + \delta} \Phi \left(\frac{q_2 - a^*}{\sigma q_1 \sqrt{a^*(1 - a^*)}} \right) \right] \in \left(0, \frac{1}{1 + \delta} \right). \end{aligned} \quad (24)$$

The upper bound follows from the fact that $a^* \in (\frac{1}{2}, 1)$, then

$$(1 - a^*) \left[1 - \frac{\delta}{1 + \delta} \Phi \left(\frac{q_2 - a^*}{\sigma q_1 \sqrt{a^*(1 - a^*)}} \right) \right] \leq \frac{1}{2} \left[1 - \frac{\delta}{1 + \delta} \frac{1}{2} \right] = \frac{2 + \delta}{4(1 + \delta)} < \frac{1}{1 + \delta},$$

for any $\delta \in [0, 1]$. Finally, suppose that $\hat{B}_{t+1} > \bar{B}_t$ ($\geq \hat{B}_t$), then we have

$$\frac{\partial Q(\mathbf{u}_t, \hat{B}_{t+1})}{\partial \hat{B}_{t+1}} = \frac{1}{1 + \delta} \left[a^*(1 + \delta) + \delta(1 - a^*) \Phi \left(\frac{1}{\sigma q_1} \sqrt{\frac{1 - a^*}{a^*}} \right) \right] \in \left(\frac{\delta}{1 + \delta}, 1 \right). \quad (25)$$

The upper bound follows from $a^* < 1$, since $q_2 = 1$ in this case, while the lower bound follows from $a^* \geq \frac{1}{2}$, by Proposition 3.

Below \bar{B}_t , \bar{Q} is flat as a function of \hat{B}_{t+1} , while the index of the new unit strictly increases in \hat{B}_{t+1} , by (23) and (24). Above \bar{B}_t , \bar{Q} grows at most by $\frac{\delta}{1 + \delta}$ following an increase in \hat{B}_{t+1} , while the derivative of the index of the new unit is at least $\frac{\delta}{1 + \delta}$ as shown by (25). If $\bar{Q}(\bar{B}_t) > Q(\mathbf{u}_t, \bar{B}_t)|_{\hat{B}_{t+1}=\bar{B}_t}$, the index of the new unit is strictly below $\bar{Q}(\bar{B}_t)$ for any $\hat{B}_{t+1} < \bar{B}_t$. Since, above \bar{B}_t , the lowest slope of the index of the new unit is strictly higher than the largest slope of $\bar{Q}(\cdot)$, the two indexes intersect exactly once. If instead $\bar{Q}(\bar{B}_t) < Q(\mathbf{u}_t, \bar{B}_t)|_{\hat{B}_{t+1}=\bar{B}_t}$, the two indexes necessarily cross only once at some $\hat{B}_{t+1} < \bar{B}_t$, but never above

³⁸Again dropping dependence on side-derivatives, see Footnote 34.

\bar{B}_t . We ignore the case $\bar{Q}(\bar{B}_t) = Q(\mathbf{u}_t, \bar{B}_t)|_{\hat{B}_{t+1}=\bar{B}_t}$, which occurs with zero probability. The intersection is denoted by b_t^I and it is such that incremental experimentation would occur, when optimal, in the newly created unit if and only if $\hat{B}_{t+1} > b_t^I$.

We now analyze how the index of radical experimentation varies with \hat{B}_{t+1} , which again denotes the new frontier payoff. With a slight abuse of notation, the index of radical experimentation at time $t+1$ is given by $Q_{t+1}^R(\hat{B}_{t+1}) = \hat{B}_{t+1} + F^\infty(G_{t+1}) = \hat{B}_{t+1} + F^\infty(\max\{\bar{B}_t - \hat{B}_{t+1}, 0\})$ as a function of the frontier payoff \hat{B}_{t+1} . For any $\hat{B}_{t+1} \geq \bar{B}_t$, $Q_{t+1}^R(\hat{B}_{t+1}) = \hat{B}_{t+1} + F^\infty(0)$ and $\frac{\partial Q_{t+1}^R}{\partial \hat{B}_{t+1}} = 1$.

Suppose instead that $\hat{B}_{t+1} < \bar{B}_t$. From Proposition 2, there exists a threshold \tilde{G} above which the optimal size of radical experimentation is 0. Define \tilde{B} by $\bar{B}_t - \tilde{B} = \tilde{G}$. Then, for any $\hat{B}_{t+1} \leq \tilde{B}$, $Q_{t+1}^R(\hat{B}_{t+1}) = \frac{\hat{B}_{t+1} + \delta \bar{B}_t}{1 + \delta}$, because the time- $t+1$ generation would prefer the frontier technology to any radical experimentation. Thus, in this case $\frac{\partial Q_{t+1}^R}{\partial \hat{B}_{t+1}} = \frac{1}{1 + \delta}$.

For $\hat{B}_{t+1} \in (\tilde{B}, \bar{B}_t)$, the maximizer of (13) may be interior. In that case, $\frac{\partial Q_{t+1}^R}{\partial \hat{B}_{t+1}} = 1 - \frac{\delta}{1 + \delta} \Phi\left(\frac{\bar{B}_t - \hat{B}_{t+1}}{\sigma \sqrt{y^*}}\right) \in \left(\frac{1}{1 + \delta}, 1\right)$, where y^* is the optimal size of radical experimentation following history \mathcal{H}_{t+1} . Thus, the derivative is always at least $\frac{1}{1 + \delta}$.

We need to consider two cases.

Case 1: $G_t = 0$. Let b_t^I denote the cutoff such that incremental experimentation occurs in the new unit if, and only if, $\hat{B}_{t+1} \geq b_t^I$, whose existence we just proved. If $b_t^I \geq \bar{B}_t$, the index of incremental experimentation is flat for any $\hat{B}_{t+1} < \bar{B}_t$, and

$$\frac{\partial Q_{t+1}^I}{\partial \hat{B}_{t+1}} \in \begin{cases} \left(0, \frac{\delta}{1 + \delta}\right) & \text{from (22) if } \bar{B}_t \leq \hat{B}_{t+1} < b_t^I, \\ \left(\frac{\delta}{1 + \delta}, 1\right) & \text{from (25) if } \hat{B}_{t+1} \geq b_t^I, \end{cases}$$

which follows from our previous analysis. Since the slope of the index of radical experimentation is always at least $\frac{1}{1 + \delta} \geq \frac{\delta}{1 + \delta}$ below b_t^I and 1 above b_t^I (as $b_t^I \geq \bar{B}_t$), it follows that Q_{t+1}^R and Q_{t+1}^I as functions of \hat{B}_{t+1} cross exactly once. Let b_t^R denote such intersection.

Similarly, if $b_t^I < \bar{B}_t$, the index for radical experimentation is unchanged, but

$$\frac{\partial Q_{t+1}^I}{\partial \hat{B}_{t+1}} \begin{cases} = 0 & \text{if } \hat{B}_{t+1} < b_t^I, \\ \in \left(0, \frac{1}{1 + \delta}\right) & \text{from (24) if } b_t^I \leq \hat{B}_{t+1} < \bar{B}_t, \\ \in \left(\frac{\delta}{1 + \delta}, 1\right) & \text{from (25) if } \hat{B}_{t+1} \geq \bar{B}_t. \end{cases}$$

A direct comparison of the slopes of the indexes shows that there exists exactly one intersection.

Case 2: $G_t > 0$. If $b_t^I \geq \bar{B}_t$, the analysis is the same as for Case 1. Thus, there exists a unique intersection b_t^R .

If $b_t^I < \bar{B}_t$, the slope of the index of incremental experimentation over the range $[\max\{b_t^I, B_l\}, \bar{B}_t]$ is given by (23), which cannot be compared with the slope of Q_{t+1}^R in an unambiguous way. Thus, we cannot exclude the possibility of multiple intersections between the two indexes. ■

B Proof of Theorem 1

Step 1: We first show that radical experimentation ends in finite time almost surely. From Proposition 2, there exists a threshold $\tilde{G} > 0$ such that radical experimentation ends after any history such that $G_t \geq \tilde{G}$. Therefore, it suffices to show that this threshold is reached almost surely.

Consider an experimentation path along which radical experimentation happens infinitely often, and let $\{\varphi(t)\}$ denote the sequence of times at which radical experimentation occurs. In particular, $y_{\varphi(t)}^R = x_{\varphi(t)}^R - \hat{x}_{\varphi(t)} > 0$ for all t .

LEMMA 2 $\{\hat{x}_{\varphi(t)}\}$ is unbounded a.s.

Proof. Suppose that $\hat{x}_{\varphi(t)}$, which is increasing in t , converges to some finite limit \tilde{x} . This implies that $y_{\varphi(t)}^R$ converges to 0. From Proposition 2, $y_{\varphi(t)}^R$ is decreasing in the gap. Therefore, there must exist a subsequence $\{\psi(t)\}$ of $\{\varphi(t)\}$ such that $\{G_{\psi(t)} = \bar{B}_{\psi(t)} - \hat{B}_{\psi(t)}\}$ is increasing. Since that sequence is bounded above by \tilde{G} , it must converge to some strictly positive limit ρ , and $\bar{B}_{\psi(t)}$ converges to the limit $\rho + B(\tilde{x})$. For sufficiently high t , the expected payoff from radical experimentation is approximately equal to $B(\tilde{x}) + \delta(\rho + B(\tilde{x}))$, while the payoff from exploitation is approximately $(1 + \delta)(\rho + B(\tilde{x}))$. Since $\rho + B(\tilde{x}) > B(\tilde{x})$ by construction, an agent will eventually strictly prefer to exploit the best explored technology yielding a payoff close to $\rho + B(\tilde{x})$ over the radical experimentation corresponding to times in $\{\varphi(t)\}$, a contradiction. ■

From Proposition 2, $\bar{y} = y^R(0)$ is an upper bound on the optimal size of radical experimentation for any size of the gap. Therefore, $|\hat{x}_{\varphi(t+1)} - \hat{x}_{\varphi(t)}| \leq \bar{y}$. For any $s > 0$ and path B , let

$$A_s(-\tilde{G}) = \sup \left\{ x' - x : \max_{x'' \in [x, x']} \{B(x'')\} < -\tilde{G}, \text{ and } x < x' < s \right\}.$$

LEMMA 3 $A_s(-\tilde{G}) > \bar{y}$ almost surely as $s \rightarrow +\infty$.

Proof. By the recurrence property of Brownian motion, there exists a.s. an $\tilde{x} > 0$ such that $B(\tilde{x}) < -\tilde{G}$ and, hence, some $s > 0$ such that $A_s(-\tilde{G}) > 0$. The result then follows from the scaling property of Brownian motion. ■

Lemma 3 means that if radical experimentation goes far enough, with each leap size bounded above by \bar{y} , the frontier is bound to “fall” into a region where its payoff is less than $-\tilde{G}$. Because \bar{B} is always nonnegative, the gap $\bar{B} - \hat{B}$ after such history will exceed \tilde{G} , prompting radical experimentation to stop. Combined with Lemma 2, this guarantees that radical experimentation must stop in finite time, almost surely.

Step 2: Belief convergence.

After radical experimentation stops, all experimentation takes place in a compact domain. The payoff distribution over that domain is characterized by finitely many Brownian bridges, whose endpoints correspond to previously explored technologies. We now establish that the beliefs resulting from the subsequent experimentation converge to a well-defined limit. Let $K = [0, \hat{x}]$ denote the domain of experimentation after radical experimentation has stopped, and μ_0 denote the distribution of B on K , given the history leading up to the end of radical experimentation. For notational simplicity, we will reset to 0 the time at which radical experimentation has stopped.

Let Θ denote the space of continuous functions on K starting at 0. At any time t the belief μ_t is a probability distribution over Θ : $\mu_t \in \Delta(\Theta)$. Some arguments that we need to use hold only for compact spaces. Because of this, we will sometimes need to replace the paths B by some bounded counterpart. For $\Lambda > 1$, we will consider any transformation $T(\cdot, \Lambda)$ of \mathbb{R} such that i) $T(\cdot, \Lambda)$ is continuous and strictly increasing, ii) $T(x, \Lambda) = x$ for all x such that $|x| < \Lambda - 1$, and iii) $\lim_{x \rightarrow -\infty} T(x, \Lambda) = -\Lambda$ and $\lim_{x \rightarrow +\infty} T(x, \Lambda) = +\Lambda$. Such function is easily built, and is bounded by Λ .

For any Brownian path B , the transformed path $b^\Lambda : x \mapsto T(B(x), \Lambda)$ is continuous and bounded by Λ , and is homeomorphic to B . In particular, b^Λ and B are observationally equivalent. Any belief μ about B has a corresponding belief μ^Λ about b^Λ and vice versa. Let $\Theta(\Lambda)$ denote the subset of Θ whose elements are bounded in absolute value by Λ , and $\Delta(\Lambda)$ denote the set of distributions over $\Theta(\Lambda)$.

Given a sequence $\{x_t\}_{t \geq 0}$ of technology choices, let $\{\mu_t^\Lambda\}$ denote the sequence of beliefs in $\Delta(\Lambda)$ about the underlying transformed path b^Λ , obtained through Bayesian updating. It is well-known that this sequence is a martingale and converges to some limit μ^Λ . This result follows from the Martingale Convergence Theorem, and is proved similarly to Theorem 4 in Easley and Kiefer (1988). Translating this result in terms of B , this shows that the sequence $\{\mu_t\}$ of beliefs about the path B also converge to some limit μ .

For any history \mathcal{H} leading to the belief μ , let $Z^\Lambda(\mu) = \sup\{b^\Lambda(x_t) : x_t \text{ contained in } \mathcal{H}\}$. As is easily checked, $Z^\Lambda(\mu)$ is independent of the particular history leading up to the limiting belief μ , and continuous in μ . We can similarly let $Z(\mu) = \max\{B(x_t)\} = T^{-1}(Z^\Lambda(\mu), \Lambda)$ ³⁹ where, for each Λ , $T^{-1}(\cdot, \Lambda)$ denotes the inverse of $T(\cdot, \Lambda)$.

Step 3: Technology Convergence and Vanishing Value of Incremental Experimentation.

The next step is to characterize the limit to which technologies converge. For any $(z_1, z_2) \in \mathbb{R} \times \mathbb{R}_+$, let $r(z_1, z_2) = z_1 + \delta \max\{z_1, z_2\}$ denote the payoff of an agent if the payoff of his chosen technology when young is z_1 and the best explored payoff until then was z_2 . Given a technology x , payoff z , and belief μ , let

$$u(x, \mu, z) = \int_{\Theta} r(B(x), z) d\mu(B),$$

and

$$u^\Lambda(x, \mu, z) = \int_{\Theta} r(T(B(x), \Lambda), z) d\mu(B).$$

Using the distribution μ^Λ implied on $\Theta(\Lambda)$ by μ , we have

$$u^\Lambda(x, \mu, z) = v(x, \mu^\Lambda, z),$$

where we define v , for any $\tilde{\mu} \in \Delta(\Lambda)$, by

$$v(x, \tilde{\mu}, z) = \int_{\Theta(\Lambda)} r(b^\Lambda(x), z) d\tilde{\mu}(b^\Lambda).$$

We will use the following lemma, which is proved at the end of this Appendix (Section B.1):

LEMMA 4 $v(x, \mu, z)$ is continuous over $K \times \Delta(\Lambda) \times [-\Lambda, \Lambda]$.

Given a belief μ with corresponding maximum explored payoff z , a young agent solves the maximization problem:

$$U(\mu, z) = \max_{x \in K} u(x, \mu, z).$$

The equilibrium technological path, denoted $\{x_t^*\}$ is such that, for each t , x_t^* maximizes $u(x, \mu_t, z_t)$. We now derive properties for the long-run technologies arising in equilibrium. Given a sequence of technologies $\{x_t\}_{t=0}^\infty$, let $\mathcal{M}(\{x_t\})$ be the set of its limit points.

PROPOSITION 10 For any history \mathcal{H} , limiting belief μ , and $x \in \mathcal{M}(\{x_t^*\})$, $x \in \operatorname{argmax}_{x' \in K} u(x', \mu, Z(\mu))$ and $B(x) = Z(\mu)$.

Proof. Let $\{x_{t_k}^*\}$ denote a subsequence converging to x . By construction,

$$u(x_{t_k}^*, \mu_{t_k}, Z(\mu_{t_k})) \geq u(x', \mu_{t_k}, Z(\mu_{t_k})), \quad (26)$$

for any $x' \in K$. Because Lemma 4 applies only to bounded payoffs, we cannot directly take the limit in the previous inequality. Instead we will approximate it by its equivalent when the payoffs are bounded by Λ for Λ arbitrarily large. Let $\Omega(\Lambda) = \{B \in \Theta : \max_{x \in K} |B(x)| > \Lambda - 1\}$. We have for any $x, \hat{\mu}, z$

$$|u(x, \hat{\mu}, z) - u^\Lambda(x, \hat{\mu}, z)| \leq \int_{\Theta} |r(B(x), z) - r(T(B(x), \Lambda), z)| d\hat{\mu}(B) \leq 2 \int_{\Omega(\Lambda)} |r(B(x), z)| d\hat{\mu}(B). \quad (27)$$

We now show that the right-hand side converges to zero as Λ goes to infinity, uniformly on the domain $K \times \cup_t \{\mu_t\} \times [0, Z(\mu)]$. For all $x \in K$ and $z \in [0, Z(\mu)]$, $|r(B(x), z)| \leq (1 + \delta)(Z(\mu) + \max_{x \in K} |B(x)|)$. Therefore, the right-hand side of (27) is bounded above by⁴⁰

$$2(1 + \delta) \int_{\Omega(\Lambda)} \left(Z(\mu) + \max_{x \in K} B(x) - \min_{x \in K} B(x) \right) d\hat{\mu}(B).$$

³⁹This maximum is also well defined, because B is continuous on the compact domain K .

⁴⁰The inequality relies on the fact that, since $B(0) = 0$, $\max_{x \in K} B(x) \geq 0$ and $\min_{x \in K} B(x) \leq 0$.

We will show that

$$\sup_{\mu_t: t \geq 0} \int_{\Omega(\Lambda)} \left(Z(\mu) + \max_{x \in K} B(x) - \min_{x \in K} B(x) \right) d\mu_t(B),$$

converges to zero as Λ goes to infinity. For this, it suffices to show the convergence for

$$\sup_{\mu_t: t \geq 0} \int_{\Omega(\Lambda)} \left(\max_{x \in K} B(x) \right) d\mu_t(B),$$

since the other two terms can be treated similarly.⁴¹ We will establish a stronger result, whose proof is in Appendix B.2. Let $\mathcal{P}(K)$ denote the set of all finite partitions of K and, for each $\Pi \in \mathcal{P}(K)$ and $\bar{Z} \geq 0$, let $\mu_{\Pi}^{\bar{Z}}$ denote the probability measure over Θ corresponding to Brownian bridges with endpoints at consecutive elements of Π and endpoint values identically equal to \bar{Z} .

LEMMA 5 *For any constant $\bar{Z} \geq 0$,*

$$\lim_{\Lambda \rightarrow +\infty} \left\{ \sup_{\Pi \in \mathcal{P}(K)} \left\{ \int_{\Omega(\Lambda)} \left(\max_{x \in K} B(x) \right) d\mu_{\Pi}^{\bar{Z}}(B) \right\} \right\} = 0.$$

For each μ_t , $\max_{x \in K} B(x)$ is clearly dominated, in the sense of first-order stochastic dominance, by the same maximum under the distribution μ_{Π} , whose partition corresponds to the units of μ_t , and whose endpoints are equal to $Z(\mu)$, which is greater than $Z(\mu_t)$.⁴² Applying Lemma 5 to $\bar{Z} = Z(\mu)$ thus proves the desired uniform convergence.

This implies that there exists, for any $\varepsilon > 0$, a positive threshold $\Lambda(\varepsilon)$ such that $|u^{\Lambda}(x, \mu_t, z) - u(x, \mu_t, z)| < \varepsilon$ for all $(x, t, z) \in K \times \mathbb{N} \times [0, Z(\mu)]$ and $\Lambda > \Lambda(\varepsilon)$. Therefore, (26) implies that, for a sequence converging to x , we have

$$\begin{aligned} u^{\Lambda}(x_{t_k}^*, \mu_{t_k}, Z(\mu_{t_k})) &\geq u(x_{t_k}^*, \mu_{t_k}, Z(\mu_{t_k})) - \varepsilon \\ &\geq u(x', \mu_{t_k}, Z(\mu_{t_k})) - \varepsilon \\ &\geq u^{\Lambda}(x', \mu_{t_k}, Z(\mu_{t_k})) - 2\varepsilon. \end{aligned}$$

Taking the limit as Λ goes to infinity, and using Lemma 4, we obtain that

$$u(x, \mu, Z(\mu)) \geq u(x', \mu, Z(\mu)) - 2\varepsilon.$$

Since ε was arbitrary, this proves that $u(x, \mu, Z(\mu)) \geq u(x', \mu, Z(\mu))$. Proposition 5 also implies that $u(x_{t_k}, \mu_{t_k}, Z(\mu_{t_k})) > (1 + \delta)Z(\mu_{t_k})$, which shows that $U(\mu, Z(\mu)) = u(x, \mu, Z(\mu)) \geq (1 + \delta)Z(\mu)$. Moreover, $B(x) \leq Z(\mu)$, since $Z(\mu) = \sup_t \{B(x_t)\}$, B is continuous, and x is a limit point of $\{x_t\}$, and $u(x, \mu, Z(\mu)) = B(x) + \delta Z(\mu) \leq (1 + \delta)Z(\mu)$, where the equality holds because $B(x)$ is known given μ . Therefore, $U(\mu, Z(\mu)) = (1 + \delta)Z(\mu)$. In particular, $B(x) = Z(\mu)$. \blacksquare

Proposition 10 and its proof also show that the value of incremental experimentation converges to zero over time: for any x , $u(x, \mu_t, Z(\mu_t)) - (1 + \delta)Z(\mu_t)$ becomes nonpositive.

⁴¹The last term obtains by symmetry, the first term with the constant $Z(\mu)$ can in fact be incorporated in \bar{Z} in the argument following Lemma 5.

⁴²Indeed, μ_{Π} is obtained from μ_t by raising to the level $Z(\mu)$ the payoff of each technology explored by time t .

B.1 Proof of Lemma 4

Let $\tilde{\Theta} = \Theta(\Lambda)$ (Λ is fixed throughout, so there is no ambiguity about the underlying space). Let $\{(x_n, \mu_n, z_n)\}$ be a sequence from $K \times \Delta(\tilde{\Theta}) \times [-\Lambda, \Lambda]$ which converges to $(x, \mu, z) \in K \times \Delta(\tilde{\Theta}) \times [-\Lambda, \Lambda]$. Then,

$$\begin{aligned}
|v(x_n, \mu_n, z_n) - v(x, \mu, z)| &= \left| \int_{\tilde{\Theta}} r(b^\Lambda(x_n), z_n) d\mu_n - \int_{\tilde{\Theta}} r(b^\Lambda(x), z) d\mu \right| \\
&\leq \left| \int_{\tilde{\Theta}} [r(b^\Lambda(x_n), z_n) - r(b^\Lambda(x_n), z)] d\mu_n \right| + \left| \int_{\tilde{\Theta}} [r(b^\Lambda(x_n), z) - r(b^\Lambda(x), z)] d\mu \right| \\
&\quad + \left| \int_{\tilde{\Theta}} [r(b^\Lambda(x_n), z) - r(b^\Lambda(x), z)] d\mu_n \right| + \left| \int_{\tilde{\Theta}} r(b^\Lambda(x), z) d\mu_n - \int_{\tilde{\Theta}} r(b^\Lambda(x), z) d\mu \right| \\
&\leq \delta |z_n - z| + 2 \int_{\tilde{\Theta}} |r(b^\Lambda(x_n), z) - r(b^\Lambda(x), z)| d\mu \\
&\quad + \int_{\tilde{\Theta}} |r(b^\Lambda(x_n), z) - r(b^\Lambda(x), z)| d\mu_n + \left| \int_{\tilde{\Theta}} r(b^\Lambda(x), z) d\mu_n - \int_{\tilde{\Theta}} r(b^\Lambda(x), z) d\mu \right|.
\end{aligned}$$

The last term converges to zero by weak convergence of the beliefs. We focus on the second term

$$\int_{\tilde{\Theta}} |r(b^\Lambda(x_n), z) - r(b^\Lambda(x), z)| d\mu \leq (1 + \delta) \int_{\tilde{\Theta}} |b^\Lambda(x_n) - b^\Lambda(x)| d\mu,$$

which converges to zero by the Bounded Convergence theorem. Next,

$$\int_{\tilde{\Theta}} |r(b^\Lambda(x_n), z) - r(b^\Lambda(x), z)| d\mu_n \leq (1 + \delta) \int_{\tilde{\Theta}} |b^\Lambda(x_n) - b^\Lambda(x)| d\mu_n.$$

K is compact and every $b^\Lambda \in \tilde{\Theta}$ is continuous, hence also uniformly continuous. Fix $\varepsilon > 0$ and let

$$A\left(\frac{1}{j}, \varepsilon\right) = \left\{ b^\Lambda \in \tilde{\Theta} : \exists \lambda > \frac{1}{j} \text{ s.t. } |x - y| < \lambda \implies |b^\Lambda(x) - b^\Lambda(y)| < \varepsilon \right\}.$$

By the previous observations, it also follows that for any $b^\Lambda \in \tilde{\Theta}$, there exists $j = j(b^\Lambda)$ such that $b^\Lambda \in A\left(\frac{1}{j}, \varepsilon\right), \forall j' > j$. Thus, $\tilde{\Theta} = \bigcup_{j=1}^{\infty} A\left(\frac{1}{j}, \varepsilon\right)$.

Next, since $\left\{ A\left(\frac{1}{j}, \varepsilon\right) \right\}$ converges to $\tilde{\Theta}$, it follows that for any $\vartheta > 0$, there exists $J > 0$ such that $\mu\left(A\left(\frac{1}{j}, \varepsilon\right)\right) > 1 - \frac{\vartheta}{2}, \forall j > J$. Fix $\tilde{j} > J$, by weak convergence of beliefs, there exists $N > 0$ such that $\left| \mu_n\left(A\left(\frac{1}{\tilde{j}}, \varepsilon\right)^c\right) - \mu\left(A\left(\frac{1}{\tilde{j}}, \varepsilon\right)^c\right) \right| < \frac{\vartheta}{2}$, for any $n > N$.

Since $x_n \rightarrow x$, there exists $N' > N$ such that $|x_n - x| < \frac{1}{\tilde{j}}$, for any $n > N'$. Finally, we obtain, for $n > N'$,

$$\begin{aligned}
\int_{\tilde{\Theta}} |b^\Lambda(x_n) - b^\Lambda(x)| d\mu_n &= \int_{A\left(\frac{1}{\tilde{j}}, \varepsilon\right)} |b^\Lambda(x_n) - b^\Lambda(x)| d\mu_n + \int_{A\left(\frac{1}{\tilde{j}}, \varepsilon\right)^c} |b^\Lambda(x_n) - b^\Lambda(x)| d\mu_n \\
&\leq \sup_{b \in A\left(\frac{1}{\tilde{j}}, \varepsilon\right)} |b(x_n) - b(x)| + 2\Lambda \mu_n\left(A\left(\frac{1}{\tilde{j}}, \varepsilon\right)^c\right) \\
&\leq \varepsilon + 2\Lambda \left[\left| \mu_n\left(A\left(\frac{1}{\tilde{j}}, \varepsilon\right)^c\right) - \mu\left(A\left(\frac{1}{\tilde{j}}, \varepsilon\right)^c\right) \right| + \mu\left(A\left(\frac{1}{\tilde{j}}, \varepsilon\right)^c\right) \right] \\
&\leq \varepsilon + 2\Lambda\vartheta.
\end{aligned}$$

Since ε and ϑ were arbitrary, this completes the proof. ■

B.2 Proof of Lemma 5

Letting $\Omega_+(\Lambda) = \{B : \max_{x \in K} B(x) > \Lambda - 1\}$ and $\Omega_-(\Lambda) = \{B : \min_{x \in K} B(x) < -(\Lambda - 1)\}$, we have $\Omega(\Lambda) = \Omega_+(\Lambda) \cup \Omega_-(\Lambda)$. Since $\bar{Z} \geq 0$, $\max_{x \in K} B(x)$ is nonnegative for all B in the support of any $\mu_\Pi \in \mathcal{P}(K)$. Therefore,

$$\int_{\Omega(\Lambda)} \max_{x \in K} B(x) d\mu_\Pi(B) \leq \int_{\Omega_+(\Lambda)} \max_{x \in K} B(x) d\mu_\Pi(B) + \int_{\Omega_-(\Lambda)} \max_{x \in K} B(x) d\mu_\Pi(B).$$

We will prove that the first term (the harder one) converges to zero as $\Lambda \rightarrow \infty$, uniformly in Π . The second term can be treated similarly.

Put in the language of probability theory, we need to show that, for each $\bar{Z} \geq 0$, the family of random variables $\{X_\Pi = \max_{x \in K} B(x) : B \sim \mu_\Pi\}_{\Pi \in \mathcal{P}(K)}$ is uniformly integrable.⁴³ To show uniform integrability, it suffices to prove that there exists $p > 1$ such that:⁴⁴

$$\sup_{\Pi \in \mathcal{P}(K)} E[X_\Pi^p] < \infty.$$

Without loss of generality, we set $\bar{Z} = 0$ (other cases follow by translation) and $K = [0, 1]$ (other cases follow by the scaling property of Brownian motion). For each Π , we have

$$Pr(X_\Pi \leq \Lambda) = \prod_{\pi_i \in \Pi} Pr(X_i \leq \Lambda),$$

where $\{\pi_i\}_i$ describes the units of the partition Π , X_i is the maximum of B over π_i , and we are using the fact that the variables $\{X_i\}_i$ are independently distributed. Moreover, each X_i is the maximum of a Brownian bridge with width η_i (the width of π) and endpoints equal to 0. This implies that⁴⁵

$$Pr(X_i \leq \Lambda) = 1 - e^{-2\Lambda^2/\eta_i^2}.$$

Therefore, we can compute the density of X_Π , and obtain, for $p = 2$,

$$E[X_\Pi^2] = \int_{\mathbb{R}_+} \Lambda^2 \sum_i \left(\prod_{j \neq i} (1 - e^{-2\Lambda^2/\eta_j^2}) \right) e^{-2\Lambda^2/\eta_i^2} \frac{4\Lambda}{\eta_i^2} d\Lambda.$$

For each i , the product with respect to j is bounded by 1, implying that

$$E[X_\Pi^2] \leq 4 \sum_i \int_{\mathbb{R}_+} \frac{\Lambda^3 e^{-2\Lambda^2/\eta_i^2}}{\eta_i^2} d\Lambda.$$

Making, for each i , the change of variable $u_i = \Lambda/\eta_i$, we obtain

$$E[X_\Pi^2] \leq 4 \sum_i \int_{\mathbb{R}_+} \eta_i u_i^3 e^{-2u_i^2} du_i.$$

Since, for any partition Π , $\sum_i \eta_i = 1$, we conclude that

$$\sup_{\Pi \in \mathcal{P}(K)} E[X_\Pi^2] \leq C,$$

where $C = 4 \int_{\mathbb{R}_+} u^3 e^{-2u^2} du < \infty$.

⁴³A family $\{X_i\}_{i \in I}$ is uniformly integrable if $\lim_{\Lambda \rightarrow +\infty} \{\sup_i \{E[|X_i| : |X_i| > \Lambda]\}\} = 0$.

⁴⁴See, e.g., Durrett (1996), Exercise 4.5.1., p. 260.

⁴⁵See, e.g., Durrett (1996), Exercise 7.8.1., p. 433. The formula given there is for a Brownian bridge with width equal to 1. The general formula obtains by the scaling property of Brownian motion, which is easily shown to be inherited by the Brownian bridge: letting $\{B_t^a\}_{t \in [0, a]}$ denote a Brownian bridge with endpoints 0 and a and endpoint values equal to 0, B_t^a has the law of $B_t - \frac{t}{a} B_a$, where B is the standard Brownian motion.

C Proof of Theorem 2

We first need to define the values of incremental and radical experimentation for the social planner. Given a discount factor δ_S and an arbitrary history \mathcal{H}_t , define the following auxiliary functions:⁴⁶

$$U^{FB,I}(\mathcal{H}_t; \delta_S) \equiv \sup_{\{x_p\}_{p=t}^{+\infty}} E_{\mathcal{H}_t} \left[B(x_t) + \sum_{p=t+1}^{+\infty} \delta_S^{p-t} (B(x_p) - c(x_p - \hat{x}_p)) \right], \quad (28)$$

s.t. $x_t \in [0, \hat{x}_t], x_p \in \mathbb{R}_+, p = t+1, \dots$

and

$$U^{FB,R}(\mathcal{H}_t; \delta_S) \equiv \sup_{\{x_p\}_{p=t}^{+\infty}} E_{\mathcal{H}_t} \left[B(x_t) - c(x_t - \hat{x}_t) + \sum_{p=t+1}^{+\infty} \delta_S^{p-t} (B(x_p) - c(x_p - \hat{x}_p)) \right]. \quad (29)$$

s.t. $x_t \geq \hat{x}_t, x_p \in \mathbb{R}_+, p = t+1, \dots$

$U^{FB,I}(\mathcal{H}_t; \delta_S)$ is the value function of a social planner with discount factor δ_S when his time- t choice is restricted to technologies that lie within the frontier associated with an arbitrary history \mathcal{H}_t . Similarly, $U^{FB,R}(\mathcal{H}_t; \delta_S)$ represents the value function when the time- t choice must be greater than or equal to the frontier. As for the baseline model, if the social planner finds optimal to exploit the best explored technology at an arbitrary time t , then he will stick to that technology forever. This is because no learning occurs from exploitation so that the social planner would start $t+1$ with exactly the same beliefs that made exploitation optimal in the previous period. Thus, despite the higher complexity of the social planner's problem, we can define the values of incremental and radical experimentation in a similar way to the baseline model up to normalization:

$$V^{FB,I}(\mathcal{H}_t; \delta_S) \equiv U^{FB,I}(\mathcal{H}_t; \delta_S) - \frac{\bar{B}_t}{1 - \delta_S}, \quad \text{and} \quad V^{FB,R}(\mathcal{H}_t; \delta_S) \equiv U^{FB,R}(\mathcal{H}_t; \delta_S) - \frac{\bar{B}_t}{1 - \delta_S}. \quad (30)$$

As the values of incremental and radical experimentation are simply an affine transformation of the value functions of the social planner's maximization problem (restricting the domains of his first choice within and outside the frontier, respectively), it also follows that the social planner prefers incremental over radical experimentation at a history \mathcal{H}_t if, and only if, $V^{FB,I}(\mathcal{H}_t; \delta_S) \geq V^{FB,R}(\mathcal{H}_t; \delta_S)$.

We start the analysis with a technical result.

LEMMA 6 *There exists a sequence $\{\tilde{G}(\delta_S)\}$ with $\tilde{G}(\delta_S) \rightarrow +\infty$ as $\delta_S \rightarrow 1$ such that the value of radical experimentation for a social planner with discount factor δ_S is positive for any gap $G \leq \tilde{G}(\delta_S)$.*

Proof. Let G and \hat{x} denote the gap and the frontier associated with an arbitrary history \mathcal{H} . Define a 1-step policy as a policy such that the social planner experiments only once and then exploits forever after. Given a history \mathcal{H} , let $V^{1,R}(G; \delta_S)$ denote the value of radical experimentation when the social planner is restricted to using only a 1-step policy at any history \mathcal{H} with associated gap G .⁴⁷ Clearly, $V^{FB,R}(\mathcal{H}; \delta_S) \geq V^{1,R}(G; \delta_S)$ as a 1-step policy is a feasible policy for the social planner. The restriction to 1-step policies has the advantage that we can explicitly write down the associated maximization problem as

$$\begin{aligned} V^{1,R}(G; \delta_S) &= \sup_{y \geq 0} \left\{ \frac{\hat{B}}{1 - \delta_S} - c(y) + \frac{\delta_S}{1 - \delta_S} \sigma \sqrt{y} \phi \left(\frac{G}{\sigma \sqrt{y}} \right) + \frac{\delta_S}{1 - \delta_S} G \Phi \left(\frac{G}{\sigma \sqrt{y}} \right) \right\} - \frac{\bar{B}}{1 - \delta_S} \\ &= \frac{1}{1 - \delta_S} \sup_{y \geq 0} \left\{ -G - (1 - \delta_S)c(y) + \delta_S \sigma \sqrt{y} \phi \left(\frac{G}{\sigma \sqrt{y}} \right) + \delta_S G \Phi \left(\frac{G}{\sigma \sqrt{y}} \right) \right\}, \end{aligned} \quad (31)$$

⁴⁶Recall the convention that $c(y) = 0$ for any $y \leq 0$.

⁴⁷When restricted to 1-step policies, the value of radical experimentation depends on the history only through the gap as in the baseline model. This is not the case for general policies which can instead depend on the history in a non-trivial way.

which is equivalent to the maximization problem for an agent in our baseline model with a discount factor equal to $\frac{\delta_S}{1-\delta_S}$. Note that, for $\delta_S = 1$, the maximization problem reduces to,⁴⁸

$$\sup_{y \geq 0} \left\{ -G + \sigma\sqrt{y}\phi\left(\frac{G}{\sigma\sqrt{y}}\right) + G\Phi\left(\frac{G}{\sigma\sqrt{y}}\right) \right\} = +\infty. \quad (32)$$

Thus, $V^{1,R}(G; 1) = +\infty$, for any $G \geq 0$. This implies that there exists a $\bar{\delta} \in (0, 1)$ such that, for any $\delta_S \geq \bar{\delta}$, we can find a positive number $\tilde{G}(\delta_S)$ such that $V^{1,R}(\tilde{G}(\delta_S); \delta_S) > 0$. By Proposition 1, it also follows that $V^{FB,R}(\mathcal{H}; \delta_S) \geq V^{1,R}(G; \delta_S) > 0$, for any $G \leq \tilde{G}(\delta_S)$.

Suppose by way of contradiction that there is no sequence $\{\tilde{G}(\delta_S)\}$ with $\delta_S > \bar{\delta}$ as defined in the proof of Lemma 6 such that the sequence diverges. Take any such sequence. Then, there must exist a $\bar{G} > 0$ such that $\tilde{G}(\delta_S) < \bar{G}$ along that sequence.⁴⁹ This implies that there also exists $G' > \bar{G}$ and a subsequence such that $V^{1,R}(G'; \delta_S) < 0$ for any δ_S along the subsequence. However, evaluating (32) at $G = G'$ implies that $V^{1,R}(G'; \delta_S)$ will eventually be positive as $\delta_S \rightarrow 1$. We thus reached a contradiction. ■

Fix a path of Brownian motion. Let $\hat{x}^{FB}(\delta_S)$ denote the frontier at which a social planner with discount factor δ_S will stop radical experimentation. Suppose by way of contradiction that $\liminf_{\delta_S \rightarrow 1} \hat{x}^{FB}(\delta_S) = \check{x} < +\infty$. Take a subsequence $\hat{x}^{FB}(\delta_{S,n})$ such that $\hat{x}^{FB}(\delta_{S,n}) \uparrow \check{x}$ and $\tilde{G}(\delta_{S,n}) \rightarrow +\infty$ as $n \rightarrow +\infty$.⁵⁰ For $\check{x} < +\infty$ to be the case, it must be that the sup norm of the path of Brownian motion over the interval $[0, \hat{x}^{FB}(\delta_{S,n})]$, defined as $\sup_{x \in [0, \hat{x}^{FB}(\delta_{S,n})]} |B(x)|$, exceeds $\frac{\tilde{G}(\delta_{S,n})}{2}$ for any n . Suppose to the contrary that there exists \bar{n} such that $\sup_{x \in [0, \hat{x}^{FB}(\delta_{S,\bar{n}})]} |B(x)| \leq \frac{\tilde{G}(\delta_{S,\bar{n}})}{2}$. By Lemma 6, the value of radical experimentation for a social planner with discount factor $\delta_{S,\bar{n}}$ would be bounded away from zero over $[0, \hat{x}^{FB}(\delta_{S,\bar{n}})]$, regardless of the history, because the associated gap can never exceed $\tilde{G}(\delta_{S,\bar{n}})$ over that interval. As the social planner keeps experimenting within the fixed interval $[0, \hat{x}^{FB}(\delta_{S,\bar{n}})]$, he will eventually explore every unit that still generates a positive value of experimentation.⁵¹ This is because the frontier stays the same during experimentation and the payoff function over $[0, \hat{x}^{FB}(\delta_{S,\bar{n}})]$ is bounded with probability 1 by the almost sure continuity of Brownian motion, as in the baseline model. Therefore, the value of incremental experimentation for the social planner must eventually converge to zero and, consequently, below the value of radical experimentation. This means that the social planner will also eventually find radical experimentation optimal, which leads to a contradiction. Hence, it must be the case that $\sup_{x \in [0, \hat{x}^{FB}(\delta_{S,n})]} |B(x)| > \frac{\tilde{G}(\delta_{S,n})}{2}$ for every n along the subsequence that we are considering.

Finally, recall that $\tilde{G}(\delta_{S,n}) \rightarrow +\infty$ as $n \rightarrow +\infty$ (or $\delta_{S,n} \rightarrow 1$). This implies that the path of Brownian motion must have an infinite sup norm over the interval $[0, \check{x}]$ but this cannot be the case because the a.s. continuity of Brownian motion implies that $\sup_{x \in [0, \check{x}]} |B(x)| < +\infty$ with probability 1. We thus reached the desired contradiction. Finally, we can conclude that $\liminf_{\delta_S \rightarrow 1} \hat{x}^{FB}(\delta_S) = +\infty$ with probability 1. ■

D Cost Externalities

In the baseline model, the path of experimentation affects incentives for radical experimentation through *i*) the expected value of new technologies, and *ii*) the opportunity cost of forgoing incremental experimentation. It is perhaps more realistic to allow radical experimentation to be directly affected by past technologies, if those technologies are helpful as inputs to produce radical experimentations. Our model may be extended along this direction, by assuming that the cost of radical experimentation is decreasing in the best available

⁴⁸The derivative of the objective function with respect to y is $\frac{\sigma}{2\sqrt{y}}\phi\left(\frac{G}{\sigma\sqrt{y}}\right) > 0$.

⁴⁹If necessary, we can restrict attention to a subsequence.

⁵⁰Such a subsequence exists by Lemma 6.

⁵¹Even if we have not explicitly defined the value of experimentation of a unit, such a concept can be easily defined following (28) and (30).

technology: the cost is given by a function $c(x - \hat{x}, \alpha \bar{B})$ for $x > \hat{x}$ with $\alpha \geq 0$ ($\alpha = 0$ corresponds to the benchmark model). We assume that, for each w , the function $c(\cdot, w)$ satisfies the same assumptions as in the baseline model.

Let $y^R(G_t, \alpha \bar{B}_t)$ denote the optimal size of radical experimentation at a history \mathcal{H}_t with associated gap G_t , and current best payoff \bar{B}_t . The following result is established similarly to Proposition 2 and is stated without proof.⁵²

PROPOSITION 11 (COMPARATIVE STATICS) *Suppose that $\alpha > 0$ and that c is submodular. Then, $y^R(G, \alpha \bar{B})$ is increasing in \bar{B} and in α .*

Monotonicity with respect to \bar{B} of the optimal size of experimentation does not necessarily imply that radical experimentation itself is fostered by a higher \bar{B} , even if we also assume that the cost $c(y, \alpha \bar{B})$ is decreasing in its second component. Indeed, when the gap is positive, a higher value of \bar{B} reduces the cost of radical experimentation, but also reduces the value of radical experimentation, as shown by Proposition 1. When the gap is zero, however (following successful radical experimentation), an increase in the best available technology always stimulates radical experimentation. Therefore, this link between cost and the best technology should result in longer waves of radical experimentation.

Monotonicity of experimentation size with respect to the gap, which was established by Proposition 2, may fail in the presence of the cost externality studied here. Indeed, the reduction in the marginal benefit of radical experimentation following a larger gap might be more than compensated by a decrease in the cost and marginal cost of radical experimentation. Without the cost externality, an increase in the gap reduces the marginal benefit while leaving the marginal cost unaffected. In that case, we already know that there is a threshold for the gap above which an agent would always set the size of radical experimentation to zero.

The next result shows that stagnation still occurs as long as the marginal cost of radical experimentation is bounded below away from zero as the best available technology becomes arbitrarily large. We now assume that c is decreasing in its second component. Let $\bar{c}(y) = \lim_{\bar{B} \rightarrow +\infty} c(y, \alpha \bar{B})$ denote the lower envelope of the cost functions $\{c(\cdot, w)\}$.⁵³

THEOREM 4 (STAGNATION) *Fix $\alpha > 0$, and suppose that $\bar{c}(\cdot)$ is increasing in a right neighborhood of $y = 0$. Then, radical experimentation ends in finite time with probability one.*

Proof. Since $\bar{c}(\cdot)$ is increasing in a right neighborhood of $y = 0$, the properties of each cost function $c(\cdot, \alpha \bar{B})$ then guarantee that $\bar{c}(\cdot)$ is increasing everywhere. Replicating the proof of Proposition 2, one can show the existence of a threshold $\tilde{G} > 0$ above which the marginal benefit of radical experimentation is strictly less than the marginal cost, at any $y > 0$, for any $\bar{B} > 0$. Thus, $y^R(G, \alpha \bar{B}) = 0$ for any $G > \tilde{G}$, regardless of the absolute level of \bar{B} .

For fixed \bar{B} , $y^R(0, \alpha \bar{B})$ is still an upper bound on the size of radical experimentation for any $G > 0$, because the incentives to perform radical experimentation are maximal with a zero gap. Moreover, $y^R(0, \alpha \bar{B})$ increases in \bar{B} , because the cost function is submodular in $(y, \alpha \bar{B})$. However, the marginal cost is bounded away from zero by the properties of the lower envelope $\bar{c}(\cdot)$, which implies that there exists $0 < \Lambda < +\infty$ such that $\lim_{\bar{B} \rightarrow +\infty} y^R(0, \alpha \bar{B}) < \Lambda$. Finally, we can now repeat the same proof as in Step 1 of Theorem 1. ■

Even if the long-run dynamics is the same with and without intergenerational cost externalities, the short-run pattern of experimentation might be significantly different in the two scenarios.

⁵²The agent's objective function is submodular in (y, α) . When there are multiple maximizers, Proposition 11 holds in the sense of the strong set order (see footnote 13).

⁵³The function $\bar{c}(\cdot)$ is well-defined because, for each $y \geq 0$, the sequence $\{c(y, w)\}$ is decreasing in w and nonnegative.

E Suboptimality of the Gittins Index Policy with Correlated Arms: an Example

Consider a model with two arms, 1 and 2, and state space $S = \{s_1, s_2, s_3\}$.⁵⁴ In state s_1 arm 1 returns 10 and arm 2 returns 5. Pulling arm 1 in state s_1 sends the world into state s_2 , in which arm 1 continues returning 10 and arm 2 returns 0. Pulling arm 2 in state s_1 sends the world into state s_3 , in which the payoff from arm 1 doubles to 20 and arm 2 continues to pay out 5. States s_2 and s_3 are absorbing states which are never left. Since arm 1 returns 10 forever when pulled in state s_1 , it has a Gittins index of 10, and the optimal payoff from pulling arm 1 first is $\frac{10}{1-\delta}$. Similarly, arm 2 has a Gittins index of 5. However, by pulling arm 2 first and arm 1 from then on, the agent can obtain a payoff of $5 + \frac{20\delta}{1-\delta}$, which is strictly better than the prescribed policy for δ sufficiently close to one.

F Patents and Radical Experimentation

One way of improving inefficiently short radical experimentation is to use patents. In our model, patents would naturally consist of transfers between consecutive generations and would be defined by exploiting the spatial nature of the technological space. In particular, one could assume that, whenever an agent undertakes radical experimentation, the set of technologies that become available at no cost, due to this radical experimentation, are patented to this agent.

Such a patent protection system can affect incentives for radical experimentation through two channels:

1. *Receiving Royalty Fees* Radical experimentation now brings a positive probability that the payoff of radical experimentation will be in some intermediate range at which further radical experimentation is suboptimal for the next generation, but incremental experimentation takes place in the new unit, as shown by Proposition 6. When this happens, a small royalty fee just paid by the new generation increases the ex ante incentives for radical experimentation. Perhaps counter-intuitively, though, even in the case of a flat royalty (let alone a more complex royalty structure), the level of that royalty affects the marginal value of radical experimentation and, therefore, the optimal size of radical experimentation. Moreover, this value may be *non-monotonic* in the patent level.⁵⁵

2. *Avoiding Royalty Fees* Precisely because an incoming generation has to pay a cost to the previous generation in order to innovate in a newly created unit, this reduces the value of incremental experimentation for that generation, relative to radical experimentation: the new generation is “pushed” towards further radical experimentation.

These incentives are substitutes of each other: the more the new generation avoids royalty fees (say, because they are fixed at high levels), and the lower the patent incentives for the old generation, and vice versa. However, regardless of which effect dominates, royalty fees, structured in this way, would foster radical experimentation.

⁵⁴This example was proposed by Philip Marx for an assignment in Strulovici’s graduate course on dynamic methods (2014).

⁵⁵This may be explained as follows: a bolder radical experimentation has a higher variance and, for given mean, a higher probability of reaching very high outcomes. When this happens, the next generation has a strong incentive to use the technological domain created by this radical experimentation, which generates royalty fees for the old generation. A higher royalty fee may therefore increase the expected marginal value of radical experimentation, spurring radical experimentation and increasing the size of radical experimentation. However, as the royalty fee gets arbitrarily large, the new generation prefers to forgo this opportunity and to innovate in older units, causing the non-monotonicity.

References

- Acemoglu, D., 2008, *Introduction to Modern Economic Growth*. Princeton University Press, Princeton, NJ.
- Aghion, P., and P. Howitt, 1992, “A Model of Growth Through Creative Destruction,” *Econometrica*, 60, 323–351.
- Akcigit, U., and Q. Liu, 2013, “The Role of Information in Innovation and Competition,” mimeo.
- Banerjee, A., 1992, “A Simple Model of Herd Behavior,” *Quarterly Journal of Economics*, 107, 797–817.
- Baumol, W. J., R. E. Litan, and C. J. Schramm, 2007, *Good Capitalism, Bad Capitalism, and the Economics of Growth and Prosperity*. Yale University Press, New Haven, CT.
- Bikhchandani, S., D. Hirshleifer, and I. Welch, 1992, “A Theory of Fads, Fashion, Custom, and Cultural Change as Information Cascades,” *Journal of Political Economy*, 100, 992–1026.
- Billingsley, P., 1968, *Convergence of Probability Measures*. John Wiley and Sons, New York, NY.
- Bolton, P., and C. Harris, 1999, “Strategic Experimentation,” *Econometrica*, 67, 349–374.
- Callander, S., 2011, “Searching and Learning by Trial-and-Error,” *American Economic Review*, 101, 2277–2308.
- Dembo, A., and O. Zeitouni, 1998, *Large Deviations Techniques and Applications*. Springer-Verlag, New York, NY.
- Durrett, R., 1996, *Probability: Theory and Examples*. Duxbury Press, 2nd ed.
- Easley, D., and N. Kiefer, 1988, “Controlling a Stochastic Process with Unknown Parameters,” *Econometrica*, 56, 1045–1064.
- Edlin, A., and C. Shannon, 1998, “Strict Monotonicity in Comparative Statics,” *Journal of Economic Theory*, 81, 201–219.
- Gittins, J., and D. Jones, 1974, “A Dynamic Allocation Index for the Sequential Allocation of Experiments,” in *Progress in Statistics*, ed. by J. Gani. North-Holland, Amsterdam, pp. 241–266.
- Gittins, J. C., 1979, “Bandit Processes and Dynamic Allocation Indices,” *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 148–177.
- Gordon, R., 2012, “Is U.S. Economic Growth Over? Faltering Innovation Confronts the Six Headwinds,” *NBER working paper series*, Working paper 18315.
- Griliches, Z., 1992, “The Search for R&D Spillovers,” *Scandinavian Journal of Economics*, 94, S29–S47.
- Hall, B., 1996, “The Private and Social Returns to Research and Development,” *Technology, R&D, and the Economy*, Brookings Institution and the American Enterprise Institute, Washington, DC, 140–83.
- Jovanovic, B., and R. Rob, 1990, “Long Waves and Short Waves: Growth Through Intensive and Extensive Search,” *Econometrica*, 58, 1391–1409.

- Karatzas, I., and S. Shreve, 1991, *Brownian Motion and Stochastic Calculus*. Springer-Verlag, New York, NY.
- Keller, G., and S. Rady, 2010, “Strategic Experimentation with Poisson Bandits,” *Theoretical Economics*, 5, 275–311.
- Keller, G., S. Rady, and M. Cripps, 2005, “Strategic Experimentation with Exponential Bandits,” *Econometrica*, 73, 39–68.
- Kortum, S., 1998, “Research, Patenting, and Technological Change,” *Econometrica*, 65, 1389–1419.
- Lee, I., 1993, “On the Convergence of Informational Cascades,” *Journal of Economic Theory*, 61, 395–411.
- Milgrom, P., and I. Segal, 2002, “Envelope Theorems for Arbitrary Choice Sets,” *Econometrica*, 2, 583–601.
- Mueller-Frank, M., and M. Pai, 2013, “Social Learning with Costly Search,” mimeo.
- Niehaus, P., 2011, “Filtered Social Learning,” *Journal of Political Economy*, 119, 686–720.
- Revuz, D., and M. Yor, 1999, *Continuous Martingales and Brownian Motion*, vol. 293. Springer Science & Business Media.
- Romer, P., 1990, “Endogenous Technological Change,” *Journal of Political Economy*, 98, 71–102.
- Ruttan, V., 2006, *Is War Necessary for Economic Growth?: Military Procurement and Technological Development*. Oxford University Press, USA.
- Slywotsky, A., 2009, “How Science Can Create Millions of New Jobs: Reigniting basic research can repair the broken U.S. business model and put Americans back to work,” *BusinessWeek*, August 27th.
- Zakaria, F., 2011, “The Future of Innovation: Can America Keep Pace?,” *Newsweek*, June 5th.

Online Appendix: Social Experimentation

Umberto Garfagnini
ITAM School of Business

Bruno Strulovici
Northwestern University

June 29, 2015

Section A of this online Appendix contains the omitted proofs of the results in Section 7 of the paper about three-period agents. Appendix B extends the baseline model to the case in which the Brownian motion has a positive drift.

We assume that differentiability holds whenever derivatives are taken. We refer to the proof of Proposition 1 and Footnote 33 in the paper for possible ways to deal with non-differentiability.

A Three-Period Agents

We first need to formally define the values of incremental and radical experimentation for an arbitrary agent. Given a discount factor δ and an arbitrary history \mathcal{H}_t , define the following auxiliary functions, for any $T = 0, 1, 2$:¹

$$U^I(\mathcal{H}_t; T) \equiv \sup_{\{x_{t+s}\}_{s=0}^{2-T}} E_{\mathcal{H}_t} \left[B(x_t) + \sum_{s=1}^{2-T} \delta^s (B(x_{t+s}) - c(x_{t+s} - \hat{x}_{t+s})) \right], \quad (1)$$

s.t. $x_t \in [0, \hat{x}_t], x_{t+s} \in \mathbb{R}_+, s = 1, \dots, 2 - T$

and

$$U^R(\mathcal{H}_t; T) \equiv \sup_{\{x_{t+s}\}_{s=0}^{2-T}} E_{\mathcal{H}_t} \left[B(x_t) - c(x_t - \hat{x}_t) + \sum_{s=1}^{2-T} \delta^s (B(x_{t+s}) - c(x_{t+s} - \hat{x}_{t+s})) \right]. \quad (2)$$

s.t. $x_t \geq \hat{x}_t, x_{t+s} \in \mathbb{R}_+, s = 1, \dots, 2 - T$

We can define the values of incremental and radical experimentation in a similar way to the

¹Recall the convention that $c(y) = 0$ for any $y \leq 0$.

baseline model up to normalization:

$$\mathcal{V}^I(\mathcal{H}_t; T) \equiv U^I(\mathcal{H}_t; T) - \sum_{s=0}^{2-T} \delta^s \bar{B}_t, \quad \text{and} \quad \mathcal{V}^R(\mathcal{H}_t; T) \equiv U^R(\mathcal{H}_t; T) - \sum_{s=0}^{2-T} \delta^s \bar{B}_t. \quad (3)$$

We focus on the problem of a newly born agent. We can write the value of radical experimentation recursively as

$$\begin{aligned} \mathcal{V}^R(\mathcal{H}_t; 0) = \sup_{x_t \geq \hat{x}_t} E_{\mathcal{H}_t} [& B(x_t) - c(x_t - \hat{x}_t) - \bar{B}_t + \delta \max\{\mathcal{V}^I(\mathcal{H}_{t+1}; 1), \mathcal{V}^R(\mathcal{H}_{t+1}; 1)\} \\ & + \delta(1 + \delta) (\max\{\bar{B}_t, B(x_t)\} - \bar{B}_t)]. \end{aligned} \quad (4)$$

A similar recursive expression holds for the value of incremental experimentation.

Lemma A.1 characterizes an important property of the intermediate expected value of experimentation, $E_{\mathcal{H}_t} [\max\{\mathcal{V}^I(\mathcal{H}_{t+1}; 1), \mathcal{V}^R(\mathcal{H}_{t+1}; 1)\} | x]$, for an arbitrary radical experiment x , which will be useful later on.

LEMMA A.1. *Fix an arbitrary history \mathcal{H}_t with associated gap $G_t > 0$, and a technology $x \geq \hat{x}_t$. Then,*

$$\frac{dE_{\mathcal{H}_t} [\max\{\mathcal{V}^I(\mathcal{H}_{t+1}; 1), \mathcal{V}^R(\mathcal{H}_{t+1}; 1)\} | x]}{d\bar{B}_t} < \frac{1 + \delta}{2}, \quad (5)$$

where \bar{B}_t denotes the highest known payoff at time t .

Proof. Let $x \geq \hat{x}_t$, then

$$\begin{aligned} & E_{\mathcal{H}_t} [\max\{\mathcal{V}^I(\mathcal{H}_{t+1}; 1), \mathcal{V}^R(\mathcal{H}_{t+1}; 1)\} | x] \\ &= \int_{\mathbb{R}} \max\{\mathcal{V}^I(\mathcal{H}_t \cup \{(x, B)\}; 1), \mathcal{V}^R(\mathcal{H}_t \cup \{(x, B)\}; 1)\} \psi(B; \hat{B}_t, \sigma^2(x - \hat{x}_t)) dB, \end{aligned} \quad (6)$$

where $\psi(B; \hat{B}_t, \sigma^2(x - \hat{x}_t))$ is the pdf of a normal distribution with mean \hat{B}_t and variance $\sigma^2(x - \hat{x}_t)$. Notice that $\mathcal{V}^R(\mathcal{H}_t \cup \{(x, B)\}; 1) = (1 + \delta)V^R(\max\{\bar{B}_t - B, 0\})$, by definition of the value of radical experimentation in the baseline model. Also,

$$\mathcal{V}^I(\mathcal{H}_t \cup \{(x, B)\}; 1) = (1 + \delta) \max\{V_{old}^I(\mathcal{H}_t \cup \{(x, B)\}), V_{new}^I(\mathcal{H}_t \cup \{(x, B)\})\}. \quad (7)$$

$V_{old}^I(\mathcal{H}_t \cup \{(x, B)\})$ is the value of incremental experimentation among the old units, which is affected by the radical experiment at time t through the possible change in the size of the highest known payoff. $V_{new}^I(\mathcal{H}_t \cup \{(x, B)\})$ is the value of incremental experimentation of the unit created

by the radical experiment x , of length $x - \hat{x}_t$, with associated left endpoint outcome \hat{B}_t and right endpoint outcome B .

Suppose that the gap is positive at time t , that is, $G_t = \bar{B}_t - \hat{B}_t > 0$, so that the technology with the highest known payoff lies strictly within the frontier \hat{x}_t at time t . We want to consider the effect of an increase in the highest known payoff at time t , \bar{B}_t . Differentiating with respect to \bar{B}_t , we obtain that

$$\begin{aligned} & \frac{dE_{\mathcal{H}_t} [\max\{\mathcal{V}^I(\mathcal{H}_{t+1}; 1), \mathcal{V}^R(\mathcal{H}_{t+1}; 1)\} | x]}{d\bar{B}_t} \\ &= \int_{-\infty}^{+\infty} \frac{d [\max\{\mathcal{V}^I(\mathcal{H}_t \cup \{(x, B)\}; 1), \mathcal{V}^R(\mathcal{H}_t \cup \{(x, B)\}; 1)\}]}{d\bar{B}_t} \psi \left(B; \hat{B}_t, \sigma^2(x - \hat{x}_t) \right) dB. \end{aligned} \quad (8)$$

First, notice that an increase in \bar{B}_t weakly reduces $\mathcal{V}^R(\mathcal{H}_t \cup \{(x, B)\}; 1)$ for any fixed B due to the monotonicity of the value of radical experimentation in the baseline model established in Proposition 1. Second, let us consider the effect on the value of incremental experimentation. Consider values of B such that $B \leq \bar{B}_t$, then the value of incremental experimentation of the old units is also decreasing in the highest known payoff from Proposition 4. Similarly, the value of incremental experimentation of the new unit is decreasing in \bar{B}_t given that we assumed that $\bar{B}_t \geq B$. Therefore,

$$\int_{-\infty}^{\bar{B}_t} \frac{d [\max\{\mathcal{V}^I(\mathcal{H}_t \cup \{(x, B)\}; 1), \mathcal{V}^R(\mathcal{H}_t \cup \{(x, B)\}; 1)\}]}{d\bar{B}_t} \psi \left(B; \hat{B}_t, \sigma^2(x - \hat{x}_t) \right) dB < 0. \quad (9)$$

Next, suppose that $B > \bar{B}_t$. Now, an increase in \bar{B}_t increases the value of experimentation of the two old units which share the technology with the highest known payoff at time t as an endpoint by Proposition 3 because $\bar{B}_{t+1} = B > \bar{B}_t$. The value of experimentation of the new unit is instead unaffected given that $B > \bar{B}_t$ renders it independent from small increases in \bar{B}_t . The same holds for the value of radical experimentation because $B > \bar{B}_t$ implies that $G_{t+1} = 0$

and thus $\mathcal{V}^R(\mathcal{H}_t \cup \{(x, B)\}; 1) = (1 + \delta)V^R(0)$. Therefore,

$$\begin{aligned}
& \int_{\bar{B}_t}^{+\infty} \frac{d [\max\{\mathcal{V}^I(\mathcal{H}_t \cup \{(x, B)\}; 1), \mathcal{V}^R(\mathcal{H}_t \cup \{(x, B)\}; 1)\}] }{d\bar{B}_t} \psi \left(B; \hat{B}_t, \sigma^2(x - \hat{x}_t) \right) dB \\
& \leq \int_{\bar{B}_t}^{+\infty} \frac{d\mathcal{V}^I(\mathcal{H}_t \cup \{(x, B)\}; 1)}{d\bar{B}_t} \psi \left(B; \hat{B}_t, \sigma^2(x - \hat{x}_t) \right) dB \\
& = \int_{\bar{B}_t}^{+\infty} (1 + \delta) \frac{dV^I(\mathcal{H}_t \cup \{(x, B)\})}{d\bar{B}_t} \psi \left(B; \hat{B}_t, \sigma^2(x - \hat{x}_t) \right) dB \\
& \leq \int_{\bar{B}_t}^{+\infty} (1 + \delta) \psi \left(B; \hat{B}_t, \sigma^2(x - \hat{x}_t) \right) dB \\
& = (1 + \delta) \left(1 - \Phi \left(\bar{B}_t; \hat{B}_t, \sigma^2(x - \hat{x}_t) \right) \right) \\
& < (1 + \delta) \left(1 - \Phi \left(\hat{B}_t; \hat{B}_t, \sigma^2(x - \hat{x}_t) \right) \right) \\
& = \frac{1 + \delta}{2}. \tag{10}
\end{aligned}$$

The first inequality follows from the previous discussion. The second inequality instead follows from equation (23) in the paper, because 1 is the upper bound on the derivative of the value of experimentation of the units that contain the previous highest known payoff, \bar{B}_t , but do not contain the current highest payoff, as $\bar{B}_t < B = \bar{B}_{t+1}$, with respect to an increase in \bar{B}_t . Finally, the third inequality follows from the fact that $\bar{B}_t > \hat{B}_t$ and the symmetry of the normal distribution. Putting (9) and (10) together yields the result. \square

We then show that the value of radical experimentation is decreasing in an agent's remaining life-span.

PROPOSITION A.1. *For any arbitrary history \mathcal{H}_t , the value of radical experimentation is (weakly) decreasing with age, that is, $\mathcal{V}^R(\mathcal{H}_t; 0) \geq \mathcal{V}^R(\mathcal{H}_t; 1) \geq \mathcal{V}^R(\mathcal{H}_t; 2) = -G_t$.*

Proof. The observation that $\mathcal{V}^R(\mathcal{H}_t; 1) \geq \mathcal{V}^R(\mathcal{H}_t; 2) = -G_t$ follows from the analysis of the baseline model. Next, we are going to show that $\mathcal{V}^R(\mathcal{H}_t; 0) \geq \mathcal{V}^R(\mathcal{H}_t; 1)$, that is, for a given history, the value of radical experimentation is weakly higher for a newly born agent than for a young agent. Note that

$$\begin{aligned}
\mathcal{V}^R(\mathcal{H}_t; 0) &= \sup_{x \geq \hat{x}_t} \left\{ -G_t - c(x - \hat{x}_t) + \delta E_{\mathcal{H}_t} \left[\max\{\mathcal{V}^I(\mathcal{H}_{t+1}; 1), \mathcal{V}^R(\mathcal{H}_{t+1}; 1)\} | x \right] \right. \\
& \quad \left. + \delta(1 + \delta) \left[-G_t + \sigma \sqrt{x - \hat{x}} \phi \left(\frac{G_t}{\sigma \sqrt{x - \hat{x}}} \right) + G \Phi \left(\frac{G_t}{\sigma \sqrt{x - \hat{x}}} \right) \right] \right\}, \tag{11}
\end{aligned}$$

$$\mathcal{V}^R(\mathcal{H}_t; 1) = \sup_{x \geq \hat{x}_t} \left\{ -c(x - \hat{x}_t) + \delta \sigma \sqrt{x - \hat{x}_t} \phi \left(\frac{G_t}{\sigma \sqrt{x - \hat{x}_t}} \right) + \delta G_t \Phi \left(\frac{G_t}{\sigma \sqrt{x - \hat{x}_t}} \right) - (1 + \delta) G_t \right\}. \quad (12)$$

Then,

$$\begin{aligned} & \mathcal{V}^R(\mathcal{H}_t; 0) - \mathcal{V}^R(\mathcal{H}_t; 1) \\ &= \sup_{x \geq \hat{x}_t} \left\{ -\delta^2 G_t - c(x - \hat{x}_t) + \delta E_{\mathcal{H}_t} \left[\max\{\mathcal{V}^I(\mathcal{H}_{t+1}; 1), \mathcal{V}^R(\mathcal{H}_{t+1}; 1)\} | x \right] \right. \\ & \quad \left. + \delta(1 + \delta) \left[\sigma \sqrt{x - \hat{x}_t} \phi \left(\frac{G_t}{\sigma \sqrt{x - \hat{x}_t}} \right) + G_t \Phi \left(\frac{G_t}{\sigma \sqrt{x - \hat{x}_t}} \right) \right] \right\} \\ & \quad - \sup_{x \geq \hat{x}_t} \left\{ -c(x - \hat{x}_t) + \delta \sigma \sqrt{x - \hat{x}_t} \phi \left(\frac{G_t}{\sigma \sqrt{x - \hat{x}_t}} \right) + \delta G_t \Phi \left(\frac{G_t}{\sigma \sqrt{x - \hat{x}_t}} \right) \right\}. \end{aligned} \quad (13)$$

Let $x^{*,1} = x^{*,1}(G_t)$ denote the optimal technology for the maximization problem associated with the value function $\mathcal{V}^R(\mathcal{H}_t; 1)$, which depends only on G_t from our baseline analysis. Then,

$$\begin{aligned} & \mathcal{V}^R(\mathcal{H}_t; 0) - \mathcal{V}^R(\mathcal{H}_t; 1) \\ & \geq -\delta^2 G_t - c(x^{*,1} - \hat{x}_t) + \delta E_{\mathcal{H}_t} \left[\max\{\mathcal{V}^I(\mathcal{H}_{t+1}; 1), \mathcal{V}^R(\mathcal{H}_{t+1}; 1)\} | x^{*,1} \right] \\ & \quad + \delta(1 + \delta) \left[\sigma \sqrt{x^{*,1} - \hat{x}_t} \phi \left(\frac{G_t}{\sigma \sqrt{x^{*,1} - \hat{x}_t}} \right) + G_t \Phi \left(\frac{G_t}{\sigma \sqrt{x^{*,1} - \hat{x}_t}} \right) \right] \\ & \quad - \left\{ -c(x^{*,1} - \hat{x}_t) + \delta \sigma \sqrt{x^{*,1} - \hat{x}_t} \phi \left(\frac{G_t}{\sigma \sqrt{x^{*,1} - \hat{x}_t}} \right) + \delta G_t \Phi \left(\frac{G_t}{\sigma \sqrt{x^{*,1} - \hat{x}_t}} \right) \right\} \\ & = \delta E_{\mathcal{H}_t} \left[\max\{\mathcal{V}^I(\mathcal{H}_{t+1}; 1), \mathcal{V}^R(\mathcal{H}_{t+1}; 1)\} | x^{*,1} \right] \\ & \quad + \delta^2 \left[\sigma \sqrt{x^{*,1} - \hat{x}_t} \phi \left(\frac{G_t}{\sigma \sqrt{x^{*,1} - \hat{x}_t}} \right) + G_t \Phi \left(\frac{G_t}{\sigma \sqrt{x^{*,1} - \hat{x}_t}} \right) - G_t \right]. \end{aligned} \quad (14)$$

Let $y^{*,1}(G) = x^{*,1}(G) - \hat{x}$ and define the function

$$\mathcal{F}(G) = \sigma \sqrt{y^{*,1}(G)} \phi \left(\frac{G}{\sigma \sqrt{y^{*,1}(G)}} \right) + G \Phi \left(\frac{G}{\sigma \sqrt{y^{*,1}(G)}} \right) - G. \quad (15)$$

Notice that $\mathcal{F}(0) = \sigma \sqrt{y^{*,1}(0)} \phi(0) > 0$, and

$$\mathcal{F}'(G) = \frac{\sigma}{2\sqrt{y^{*,1}(G)}} \phi \left(\frac{G}{\sigma \sqrt{y^{*,1}(G)}} \right) \frac{dy^{*,1}(G)}{dG} + \Phi \left(\frac{G}{\sigma \sqrt{y^{*,1}(G)}} \right) - 1 < 0, \quad (16)$$

because $\frac{dy^{*,1}(G)}{dG} < 0$ by Proposition 2 in the main text. Therefore,

$$\inf_{G \geq 0} \mathcal{F}(G) = \lim_{G \uparrow +\infty} \mathcal{F}(G) = \lim_{G \uparrow +\infty} \sigma \sqrt{y^{*,1}(G)} \phi \left(\frac{G}{\sigma \sqrt{y^{*,1}(G)}} \right) - \lim_{G \uparrow +\infty} G \left(1 - \Phi \left(\frac{G}{\sigma \sqrt{y^{*,1}(G)}} \right) \right). \quad (17)$$

The first limit in (17) equals zero due to the properties of the standard normal pdf and the fact that $y^{*,1}(G) \in [0, y^{*,1}(0)]$ for any $G \geq 0$.

By applying L'Hôpital's rule for an arbitrary $y > 0$, we obtain that

$$\lim_{G \rightarrow +\infty} G \left(1 - \Phi \left(\frac{G}{\sigma \sqrt{y}} \right) \right) = \lim_{G \rightarrow +\infty} \frac{\phi \left(\frac{G}{\sigma \sqrt{y}} \right)}{G^{-2}} = 0. \quad (18)$$

As $y^{*,1}(G)$ is bounded and monotonically decreasing in G , it also follows that

$$0 \leq G \left(1 - \Phi \left(\frac{G}{\sigma \sqrt{y^{*,1}(G)}} \right) \right) \leq G \left(1 - \Phi \left(\frac{G}{\sigma \sqrt{y^{*,1}(0)}} \right) \right). \quad (19)$$

Taking limits gives that $\inf_{G \geq 0} \mathcal{F}(G) = 0$. This shows that $\mathcal{V}^R(\mathcal{H}_t; 0) - \mathcal{V}^R(\mathcal{H}_t; 1) \geq 0$ given that

$$E_{\mathcal{H}_t} [\max\{\mathcal{V}^I(\mathcal{H}_{t+1}; 1), \mathcal{V}^R(\mathcal{H}_{t+1}; 1)\} | x^{*,1}] \geq E_{\mathcal{H}_t} [\mathcal{V}^I(\mathcal{H}_{t+1}; 1) | x^{*,1}] \geq 0, \quad (20)$$

by Proposition 5. This completes the proof. \square

We can now show the following result which generalizes the idea that a larger gap reduces the value of radical experimentation even for a longer life span. As in the baseline model, we also show that a negative value of radical experimentation leads to the abandonment of that type of experimentation.

PROPOSITION A.2 (VALUE OF RADICAL EXPERIMENTATION). *(i) The value of radical experimentation for a newly born agent, $\mathcal{V}^R(\mathcal{H}_t; 0)$, is decreasing in the size of the highest known payoff, \bar{B}_t , all else equal. (ii) If the value of radical experimentation is negative at some time t , radical experimentation is abandoned forever after.*

Proof. (i) Differentiating $\mathcal{V}^R(\mathcal{H}_t; 0)$ with respect to \bar{B}_t , while keeping everything else fixed, and applying the envelope theorem gives

$$-1 - \delta(1 + \delta) \left[1 - \Phi \left(\frac{G_t}{\sigma \sqrt{x^{*,0} - \hat{x}}} \right) \right] + \delta \frac{dE_{\mathcal{H}_t} [\max\{\mathcal{V}^I(\mathcal{H}_{t+1}; 1), \mathcal{V}^R(\mathcal{H}_{t+1}; 1)\} | x^{*,0}]}{d\bar{B}_t}, \quad (21)$$

where $x^{*,0}$ is a maximizer. Lemma A.1 implies that (21) is negative. Therefore, it follows that the value of radical experimentation is decreasing in the size of the highest known payoff.

(ii) Fix a history \mathcal{H}_t . Recall that $\mathcal{V}^R(\mathcal{H}_t; 1) = (1 + \delta)V^R(G_t)$ where G_t is the gap induced by history \mathcal{H}_t . The monotonicity of $V^R(G)$ with respect to the size of the gap G , which follows from Proposition 1 in the main text, implies that, for any given technology x lying inside the current frontier and any associated outcome $B(x)$, $\mathcal{V}^R(\mathcal{H}_t; 1) \geq \mathcal{V}^R(\mathcal{H}_{t+1}; 1)$ where $\mathcal{H}_{t+1} = \mathcal{H}_t \cup \{(x, B(x))\}$ because $\max\{\bar{B}_t, B(x)\} \geq \bar{B}_t$.

Next, notice that $\mathcal{V}^R(\mathcal{H}_t; 0) < 0$ implies that the newly born agent at time t will choose either exploitation or incremental experimentation in his first period. Let x_t^* denote the technology thus chosen and let $B(x_t^*)$ denote the associated outcome. Then, the observation made in the previous paragraph and Proposition A.1 imply that $\mathcal{V}^R(\mathcal{H}_{t+1}; 1) \leq \mathcal{V}^R(\mathcal{H}_t; 1) \leq \mathcal{V}^R(\mathcal{H}_t; 0) < 0$, where $\mathcal{H}_{t+1} = \mathcal{H}_t \cup \{(x_t^*, B(x_t^*))\}$. Therefore, the newly born agent will also find radical experimentation suboptimal when young.

Finally, let us consider the incentives to perform radical experimentation of the subsequent newly born generation at time $t + 2$. We know from our previous analysis that no radical experimentation has taken place at times t and $t + 1$. Clearly, if exploitation has taken place at both t and $t + 1$, then $\mathcal{V}^R(\mathcal{H}_{t+2}; 0) = \mathcal{V}^R(\mathcal{H}_t; 0) < 0$. However, incremental experimentation might have taken place at either or even both times. As incremental experimentation weakly increases the highest known payoff, the incentives to perform radical experimentation can only decrease. \square

Next, let $x_t^{R,0}(\mathcal{H}_t)$ denote the technology which maximizes the value of radical experimentation for a newly born agent.

PROPOSITION A.3. *If $G_t = 0$, then $x_t^{R,0}(\mathcal{H}_t) > \hat{x}_t$.*

Proof. The first order condition of the maximization problem of a newly born agent pursuing radical experimentation is given by

$$\frac{\delta(1 + \delta)\sigma}{2\sqrt{x - \hat{x}_t}}\phi\left(\frac{G_t}{\sigma\sqrt{x - \hat{x}_t}}\right) + \delta\frac{dE_{\mathcal{H}_t}[\max\{\mathcal{V}^I(\mathcal{H}_{t+1}; 1), \mathcal{V}^R(\mathcal{H}_{t+1}; 1)\}|x]}{dx} = c'(x - \hat{x}_t). \quad (22)$$

Assuming differentiability once again, we also have that

$$\begin{aligned}
& \frac{dE_{\mathcal{H}_t} [\max\{\mathcal{V}^I(\mathcal{H}_{t+1}; 1), \mathcal{V}^R(\mathcal{H}_{t+1}; 1)\} | x]}{dx} = \\
& = \int_{\mathbb{R}} \frac{d [\max\{\mathcal{V}^I(\mathcal{H}_t \cup \{(x, B)\}; 1), \mathcal{V}^R(\mathcal{H}_t \cup \{(x, B)\}; 1)\}]}{dx} \psi \left(B; \hat{B}_t, \sigma^2(x - \hat{x}_t) \right) dB \\
& + \int_{\mathbb{R}} \max\{\mathcal{V}^I(\mathcal{H}_t \cup \{(x, B)\}; 1), \mathcal{V}^R(\mathcal{H}_t \cup \{(x, B)\}; 1)\} \frac{d\psi \left(B; \hat{B}_t, \sigma^2(x - \hat{x}_t) \right)}{dx} dB. \quad (23)
\end{aligned}$$

Fix $B \in \mathbb{R}$ and let us consider the effect of a small increase in the size of the radical experiment on $\mathcal{V}^R(\mathcal{H}_t \cup \{(x, B)\}; 1)$. Such an increase has no effect on the value of radical experimentation starting from $t + 1$ because the new value of radical experimentation depends only on the size of the gap which, fixing B , is independent from x . Thus, $\frac{d\mathcal{V}^R(\mathcal{H}_t \cup \{(x, B)\}; 1)}{dx} = 0$. Next, consider the effect of a similar increase on the value of experimentation of the old units, such an effect is also zero because the radical experiment at time t affects the value of experimentation of the old units at time $t + 1$ only through the outcome of the experiment, not the size of the experiment. Therefore, $\frac{d\mathcal{V}_{old}^I(\mathcal{H}_t \cup \{(x, B)\}; 1)}{dx} = 0$. However, the size of the radical experiment at time t has a direct effect on the value of experimentation of the newly created unit because of the increase in its length. From the baseline model, we know that $\frac{d\mathcal{V}_{new}^I(\mathcal{H}_t \cup \{(x, B)\}; 1)}{dx} > 0$. This shows that the first integral in (23) is non-negative.

Next,

$$\frac{d\psi \left(B; \hat{B}_t, \sigma^2(x - \hat{x}_t) \right)}{dx} = \left[-1 + \frac{(B - \hat{B}_t)^2}{\sigma^2(x - \hat{x}_t)} \right] \frac{\psi \left(B; \hat{B}_t, \sigma^2(x - \hat{x}_t) \right)}{2(x - \hat{x}_t)}. \quad (24)$$

which is also positive for radical experiments close to the frontier. Therefore, when $G_t = 0$, the size of radical experimentation is strictly positive given that the first term in (22) is unbounded close to the frontier. \square

PROPOSITION A.4. *There exists $\bar{Y} > 0$ such that $x_t^{R,0}(\mathcal{H}_t) - \hat{x}_t < \bar{Y}$ for any history \mathcal{H}_t .*

Proof. Suppose by way of contradiction that the claim is false. Take a monotone sequence $\{\bar{Y}_k\}$ with $\lim_{k \rightarrow +\infty} \bar{Y}_k = +\infty$. Then, for any k , there must exist a history \mathcal{H}_k such that $y_k^{R,0} \equiv x^{R,0}(\mathcal{H}_k) - \hat{x}_k > \bar{Y}_k$, and without loss of generality we can also assume that the sequence $\{y_k^{R,0}\}$ is monotone increasing. From the first-order condition (22), it follows that the left-hand side converges to $+\infty$ along the sequence $\{y_k^{R,0}\}$ by the strict convexity of the cost function, while

the first term on the right-hand side converges to zero. Therefore, it must also be the case that

$$\lim_{k \rightarrow +\infty} \frac{dE_{\mathcal{H}_k} [\max\{\mathcal{V}^I(\mathcal{H}_{k,+1}; 1), \mathcal{V}^R(\mathcal{H}_{k,+1}; 1)\} | x]}{dx} \Big|_{x=x^{R,0}(\mathcal{H}_k)} = +\infty, \quad (25)$$

where $\mathcal{H}_{k,+1} \equiv \mathcal{H}_k \cup (x^{R,0}(\mathcal{H}_k), B(x^{R,0}(\mathcal{H}_k)))$.

First, suppose that $\max\{\mathcal{V}^I(\mathcal{H}_{k,+1}; 1), \mathcal{V}^R(\mathcal{H}_{k,+1}; 1)\} = \mathcal{V}_{old}^I(\mathcal{H}_{k,+1}; 1)$ which is the value of incremental experimentation over the old units. Notice that $\frac{dE_{\mathcal{H}_k} [\mathcal{V}_{old}^I(\mathcal{H}_{k,+1}; 1) | x]}{dx} \Big|_{x=x^{R,0}(\mathcal{H}_k)}$ is necessarily uniformly bounded because the radical experiment affects the value of incremental experimentation over the old units only through the outcome at the new frontier, not through the size of the experiment. Thus, to complete the proof, it suffices to show that whenever experimentation is either radical or incremental over the new unit, the derivative of the expectation is still bounded above which would deliver the desired contradiction.

We can write the derivative of the expectation as

$$\begin{aligned} & \frac{dE_{\mathcal{H}_t} [\max\{\mathcal{V}_{new}^I(\mathcal{H}_{k,+1}; 1), \mathcal{V}^R(\mathcal{H}_{k,+1}; 1)\} | x^{R,0}(\mathcal{H}_k)]}{dx} = \\ & = \int_{\mathbb{R}} \frac{d \left[\max\{\mathcal{V}_{new}^I(\mathcal{H}_k \cup \{(x_k^{R,0}, B)\}; 1), \mathcal{V}^R(\mathcal{H}_k \cup \{(x_k^{R,0}, B)\}; 1)\} \right]}{dx} \psi \left(B; \hat{B}_k, \sigma^2 y_k^{R,0} \right) dB \end{aligned} \quad (26)$$

$$+ \int_{\mathbb{R}} \max\{\mathcal{V}_{new}^I(\mathcal{H}_k \cup \{(x_k^{R,0}, B)\}; 1), \mathcal{V}^R(\mathcal{H}_k \cup \{(x_k^{R,0}, B)\}; 1)\} \frac{d\psi \left(B; \hat{B}_k, \sigma^2 y_k^{R,0} \right)}{dx} dB. \quad (27)$$

Let us focus on the integral (26),

$$\begin{aligned}
& \int_{\mathbb{R}} \frac{d \left[\max\{\mathcal{V}_{new}^I(\mathcal{H}_k \cup \{(x_k^{R,0}, B)\}; 1), \mathcal{V}^R(\mathcal{H}_k \cup \{(x_k^{R,0}, B)\}; 1)\} \right]}{dx} \psi \left(B; \hat{B}_k, \sigma^2 y_k^{R,0} \right) dB \\
& \leq \int_{\mathbb{R}} \frac{d\mathcal{V}_{new}^I(\mathcal{H}_k \cup \{(x_k^{R,0}, B)\})}{dx} \psi \left(B; \hat{B}_k, \sigma^2 y_k^{R,0} \right) dB \\
& = \int_{\mathbb{R}} (1 + \delta) \frac{d\mathcal{V}_{new}^I(\mathcal{H}_k \cup \{(x_k^{R,0}, B)\}; 1)}{dx} \psi \left(B; \hat{B}_k, \sigma^2 y_k^{R,0} \right) dB \\
& \leq \int_{\mathbb{R}} \frac{\delta \sigma \phi(0)}{4 \sqrt{y_k^{R,0}}} \psi \left(B; \hat{B}_k, \sigma^2 y_k^{R,0} \right) dB \\
& = \frac{\delta \sigma \phi(0)}{4 \sqrt{y_k^{R,0}}}. \tag{28}
\end{aligned}$$

The first inequality follows from the discussion in the proof of Proposition A.3, while the second inequality follows from (19) in the main text. Therefore, as k goes to $+\infty$, the limit superior of (26) is bounded above by 0.

Next, consider the integral (27). Let $\mathcal{A}_k \subseteq \mathbb{R}$ denote the set over which $\frac{d\psi(B; \hat{B}_k, \sigma^2 y_k^{R,0})}{dx} > 0$, that is, $\mathcal{A}_k = \left\{ B \in \mathbb{R} : |B - \hat{B}_k| > \sigma \sqrt{y_k^{R,0}} \right\}$. Then, we can write the integral in (27) as

$$\begin{aligned}
& \int_{\mathbb{R}} \max\{\mathcal{V}_{new}^I(\mathcal{H}_k \cup \{(x_k^{R,0}, B)\}; 1), \mathcal{V}^R(\mathcal{H}_k \cup \{(x_k^{R,0}, B)\}; 1)\} \frac{d\psi \left(B; \hat{B}_k, \sigma^2 y_k^{R,0} \right)}{dx} dB \\
& = \underbrace{\int_{\mathbb{R} \setminus \mathcal{A}_k} \max\{\mathcal{V}_{new}^I(\mathcal{H}_k \cup \{(x_k^{R,0}, B)\}; 1), \mathcal{V}^R(\mathcal{H}_k \cup \{(x_k^{R,0}, B)\}; 1)\} \frac{d\psi \left(B; \hat{B}_k, \sigma^2 y_k^{R,0} \right)}{dx} dB}_{\leq 0} \\
& + \int_{\mathcal{A}_k} \max\{\mathcal{V}_{new}^I(\mathcal{H}_k \cup \{(x_k^{R,0}, B)\}; 1), \mathcal{V}^R(\mathcal{H}_k \cup \{(x_k^{R,0}, B)\}; 1)\} \frac{d\psi \left(B; \hat{B}_k, \sigma^2 y_k^{R,0} \right)}{dx} dB. \tag{29}
\end{aligned}$$

Then,

$$\begin{aligned}
& \int_{\mathcal{A}_k} \max\{\mathcal{V}_{new}^I(\mathcal{H}_k \cup \{(x_k^{R,0}, B)\}; 1), \mathcal{V}^R(\mathcal{H}_k \cup \{(x_k^{R,0}, B)\}; 1)\} \frac{d\psi(B; \hat{B}_k, \sigma^2 y_k^{R,0})}{dx} dB \\
&= - \underbrace{\int_{\mathcal{A}_k} \max\{\mathcal{V}_{new}^I(\mathcal{H}_k \cup \{(x_k^{R,0}, B)\}; 1), \mathcal{V}^R(\mathcal{H}_k \cup \{(x_k^{R,0}, B)\}; 1)\} \frac{\psi(B; \hat{B}_k, \sigma^2 y_k^{R,0})}{2y_k^{R,0}} dB}_{\leq 0} \\
&+ \int_{\mathcal{A}_k} \max\{\mathcal{V}_{new}^I(\mathcal{H}_k \cup \{(x_k^{R,0}, B)\}; 1), \mathcal{V}^R(\mathcal{H}_k \cup \{(x_k^{R,0}, B)\}; 1)\} (B - \hat{B}_k)^2 \frac{\psi(B; \hat{B}_k, \sigma^2 y_k^{R,0})}{2\sigma^2 (y_k^{R,0})^2} dB \\
&\leq 2 \int_{\mathcal{A}_k^+} \max\{\mathcal{V}_{new}^I(\mathcal{H}_k \cup \{(x_k^{R,0}, B)\}; 1), \mathcal{V}^R(\mathcal{H}_k \cup \{(x_k^{R,0}, B)\}; 1)\} (B - \hat{B}_k)^2 \frac{\psi(B; \hat{B}_k, \sigma^2 y_k^{R,0})}{2\sigma^2 (y_k^{R,0})^2} dB,
\end{aligned} \tag{30}$$

where $\mathcal{A}_k^+ = \left\{ B \in \mathbb{R} : B > \hat{B}_k + \sigma \sqrt{y_k^{R,0}} \text{ and } B - \hat{B}_k > 0 \right\}$. The last inequality follows from the observation that the maximum of the two values of experimentation is larger when $B > \hat{B}_k$, for any k along the sequence.

Let us further partition \mathcal{A}_k^+ into the two subsets $\mathcal{A}_k^{+,IN}$ and $\mathcal{A}_k^{+,R}$ such that

$$\begin{aligned}
& \int_{\mathcal{A}_k^+} \max\{\mathcal{V}_{new}^I(\mathcal{H}_k \cup \{(x_k^{R,0}, B)\}; 1), \mathcal{V}^R(\mathcal{H}_k \cup \{(x_k^{R,0}, B)\}; 1)\} (B - \hat{B}_k)^2 \frac{\psi(B; \hat{B}_k, \sigma^2 y_k^{R,0})}{2\sigma^2 (y_k^{R,0})^2} dB \\
&= \int_{\mathcal{A}_k^{+,IN}} \mathcal{V}_{new}^I(\mathcal{H}_k \cup \{(x_k^{R,0}, B)\}; 1) (B - \hat{B}_k)^2 \frac{\psi(B; \hat{B}_k, \sigma^2 y_k^{R,0})}{2\sigma^2 (y_k^{R,0})^2} dB \\
&+ \int_{\mathcal{A}_k^{+,R}} \mathcal{V}^R(\mathcal{H}_k \cup \{(x_k^{R,0}, B)\}; 1) (B - \hat{B}_k)^2 \frac{\psi(B; \hat{B}_k, \sigma^2 y_k^{R,0})}{2\sigma^2 (y_k^{R,0})^2} dB.
\end{aligned} \tag{31}$$

Next, notice that, for any $B \in \mathcal{A}_k^{+,IN}$,

$$\begin{aligned}
& \mathcal{V}_{new}^I(\mathcal{H}_k \cup \{(x_k^{R,0}, B)\}; 1) \\
&= -(1 + \delta)(\bar{B}_k - \hat{B}_k) + \max_{a \in [0,1]} \left\{ (B - \hat{B}_k)a \left(1 + \delta - \delta \Phi \left(\frac{(\bar{B}_k - \hat{B}_k) - (B - \hat{B}_k)a}{\sigma \sqrt{a(1-a)y_k^{R,0}}} \right) \right) \right. \\
&\quad + \delta \sigma \sqrt{a(1-a)} \sqrt{y_k^{R,0}} \phi \left(\frac{(\bar{B}_k - \hat{B}_k) - (B - \hat{B}_k)a}{\sigma \sqrt{a(1-a)y_k^{R,0}}} \right) \\
&\quad \left. + \delta (\bar{B}_k - \hat{B}_k) \Phi \left(\frac{(\bar{B}_k - \hat{B}_k) - (B - \hat{B}_k)a}{\sigma \sqrt{a(1-a)y_k^{R,0}}} \right) \right\}, \tag{32}
\end{aligned}$$

from (14) in the main text. Over the set $\mathcal{A}_k^{+,IN}$, an upper bound for the value of incremental experimentation of the new unit is provided by a similar unit in which the payoffs at both endpoints are all equal to the largest known payoff, that is, $B^L = B^H = \bar{B}_k$ if $B \leq \bar{B}_k$ and $B^L = B^H = B$ if $B > \bar{B}_k$. It can be shown from the analysis of our baseline model that the optimal technology in a bounded unit with equal payoffs at the endpoints is given by the mid-point technology. Putting all these observations together shows that, for any $B \in \mathcal{A}_k^{+,IN}$,

$$\mathcal{V}_{new}^I(\mathcal{H}_k \cup \{(x_k^{R,0}, B)\}; 1) \leq \frac{\sigma \delta}{2} \sqrt{y_k^{R,0}} \phi(0). \tag{33}$$

The value of radical experimentation for a young agent is instead uniformly bounded by $(1 +$

$\delta)V^R(0)$ regardless of the history. Thus,

$$\begin{aligned}
& \int_{\mathcal{A}_k^{+,IN}} \mathcal{V}_{new}^I(\mathcal{H}_k \cup \{(x_k^{R,0}, B)\}; 1)(B - \hat{B}_k)^2 \frac{\psi(B; \hat{B}_k, \sigma^2 y_k^{R,0})}{2\sigma^2 (y_k^{R,0})^2} dB \\
& + \int_{\mathcal{A}_k^{+,R}} \mathcal{V}^R(\mathcal{H}_k \cup \{(x_k^{R,0}, B)\}; 1)(B - \hat{B}_k)^2 \frac{\psi(B; \hat{B}_k, \sigma^2 y_k^{R,0})}{2\sigma^2 (y_k^{R,0})^2} dB \\
& \leq \int_{\mathcal{A}_k^{+,IN}} \frac{\sigma\delta}{2} \sqrt{y_k^{R,0}} \phi(0)(B - \hat{B}_k)^2 \frac{\psi(B; \hat{B}_k, \sigma^2 y_k^{R,0})}{2\sigma^2 (y_k^{R,0})^2} dB \\
& + \int_{\mathcal{A}_k^{+,R}} (1 + \delta)V^R(0)(B - \hat{B}_k)^2 \frac{\psi(B; \hat{B}_k, \sigma^2 y_k^{R,0})}{2\sigma^2 (y_k^{R,0})^2} dB \\
& \leq \left[\frac{\delta\phi(0)}{4\sigma(y_k^{R,0})^{3/2}} + (1 + \delta) \frac{V^R(0)}{2\sigma^2 (y_k^{R,0})^2} \right] \int_{\mathbb{R}} (B - \hat{B}_k)^2 \psi(B; \hat{B}_k, \sigma^2 y_k^{R,0}) dB \\
& = \left[\frac{\delta\phi(0)}{4\sigma(y_k^{R,0})^{3/2}} + (1 + \delta) \frac{V^R(0)}{2\sigma^2 (y_k^{R,0})^2} \right] \sigma^2 y_k^{R,0} \\
& = \frac{\delta\sigma\phi(0)}{4\sqrt{y_k^{R,0}}} + (1 + \delta) \frac{V^R(0)}{2y_k^{R,0}}, \tag{34}
\end{aligned}$$

where the second inequality follows from the increase in the range of integration and the non-negativity of both integrands, while the first equality follows from the definition of variance and the normal distribution. The expression in (34) converges to zero as k increases.

We have thus shown that the limit superior in (25) is bounded above, which gives a contradiction. \square

PROPOSITION A.5. *There exists a value of the gap \hat{G} such that $\mathcal{V}^R(\mathcal{H}_t; 0) < 0$ whenever $G_t > \hat{G}$.*

Proof. Let $x^{R,0}(\mathcal{H})$ denote the optimal size of radical innovation for a newly born agent at an arbitrary history \mathcal{H} . Taking the limit with respect to the size of the highest known payoff \bar{B}

gives

$$\begin{aligned}
\limsup_{\bar{B} \rightarrow +\infty} \mathcal{V}^R(\mathcal{H}; 0) &\leq \limsup_{\bar{B} \rightarrow +\infty} \left\{ -c(x_t^{R,0}(\mathcal{H}) - \hat{x}) \right\} \\
&+ \limsup_{\bar{B} \rightarrow +\infty} \left\{ -G(\bar{B}) + \delta E_{\mathcal{H}} \left[\max\{\mathcal{V}^I(\mathcal{H}_{+1}; 1), \mathcal{V}^R(\mathcal{H}_{+1}; 1)\} |x^{R,0}(\mathcal{H})\right] \right\} \\
&+ \limsup_{\bar{B} \rightarrow +\infty} \left\{ -\delta(1 + \delta)G(\bar{B}) \left(1 - \Phi \left(\frac{G(\bar{B})}{\sigma \sqrt{x^{R,0}(\mathcal{H}) - \hat{x}}} \right) \right) \right\} \\
&+ \limsup_{\bar{B} \rightarrow +\infty} \delta(1 + \delta)\sigma \sqrt{x^{R,0}(\mathcal{H}) - \hat{x}} \phi \left(\frac{G(\bar{B})}{\sigma \sqrt{x^{R,0}(\mathcal{H}) - \hat{x}}} \right),
\end{aligned} \tag{35}$$

where $G(\bar{B}) = \bar{B} - \hat{B}$. By Proposition A.4, $x^{R,0}(\mathcal{H}) - \hat{x} \in (0, \bar{Y})$, for any arbitrary history \mathcal{H} . Therefore, the first limit in (35) is nonnegative. The second limit is $-\infty$ given that the expectation grows in \bar{B} at a rate bounded above uniformly by $\frac{1+\delta}{2}$, by Lemma A.1, while the third and fourth limits are both zero. Thus, there exists a large enough gap \hat{G} such that the value of radical experimentation $\mathcal{V}^R(\mathcal{H}; 0)$ is negative for any size of the gap above \hat{G} . \square

Putting together all the results obtained so far allows us to replicate the same steps as in the proof of Theorem 1 in the main text. We can thus state the following result.

THEOREM A.1. *Consider the extended model with three-period-lived agents. Radical experimentation ends in finite time with probability one.*

B Positive Drift: Optimistic Beliefs and Stagnation

Theorem 1 in the main text assumed that the drift κ which determines the expected payoff of new technologies, was equal to zero. What happens if agents are more optimistic about κ ? Could that be enough to sustain radical experimentation in the long run?

For this problem to have a well-defined solution, we must assume that the incremental cost of radical experimentation eventually exceeds the drift κ . Otherwise, young agents may want to choose an infinite amount of radical experimentation. Precisely, we assume that $\lim_{y \rightarrow +\infty} c'(y) > \kappa(1 + \delta)$. For technical reasons, we also assume that $c''(0) > 0$.

The scenario of a positive drift creates an incentive for old agents to innovate as well, as captured by the objective function

$$U^{O,R}(\mathcal{H}_t) = \max_{x \in [\hat{x}_t, +\infty)} E_{\mathcal{H}_t} [B(x) - c(x - \hat{x}_t)] = \hat{B}_t + \kappa(x - \hat{x}_t) - c(x - \hat{x}_t). \tag{36}$$

This slightly complicates the short-run equilibrium behavior of experimentation, because agents now learn from both old and young generations. This does not, however, change the insight of Theorem 1.

THEOREM B.1. *Radical experimentation stops in finite time, almost surely.*

This extension of Theorem 1 is not completely straightforward, and may be explained as follows. The proof can be found at the end of the section.

The size of radical experimentation by the old generation, whenever it takes place, is characterized by the first-order condition

$$\kappa = c'(x - \hat{x}_t). \quad (37)$$

Let $y^{O,R}$ denote the solution to (37) and $\xi = \kappa y^{O,R} - c(y^{O,R}) > 0$. An old agent prefers radical experimentation over exploitation if and only if

$$\hat{B}_t + \kappa y^{O,R} - c(y^{O,R}) \geq \bar{B}_t, \text{ or, equivalently, if } G_t \leq \xi. \quad (38)$$

When choosing the size of his experimentation, a young agent considers the effect of his action today on his incentives tomorrow. This effect can be quantified in a perceived reduction of the gap from G_t to $G_t - \xi$ due to the possibility of performing radical experimentation in the second period, which results in higher incentives to perform radical experimentation today. However as the gap increases, ξ becomes negligible and eventually the relative benefit of radical experimentation over exploitation falls short of the explicit cost of experimentation. Thus the young agent will eventually opt for incremental experimentation, which is still strictly better than exploitation.

Proof of Theorem B.1. We start by showing the existence of a threshold \tilde{G} above which $y^R(G) = 0$.² Suppose, first, that $\kappa \leq c'(0)$, so that an old agent chooses the technology with the highest known payoff (as observed in the main text). The expected utility of a young agent from choosing technology $x > \hat{x}$ is

$$\hat{B} + \kappa y - c(y) + \delta \left\{ (\hat{B} + \kappa y) \left(1 - \Phi \left(\frac{G - \kappa y}{\sigma \sqrt{y}} \right) \right) + \sigma \sqrt{y} \phi \left(\frac{G - \kappa y}{\sigma \sqrt{y}} \right) \right\} \quad (39)$$

$$+ (G + \hat{B}) \Phi \left(\frac{G - \kappa y}{\sigma \sqrt{y}} \right) \quad (40)$$

$$= (1 + \delta)[\hat{B} + \kappa y] - c(y) + \delta \left\{ (G - \kappa y) \Phi \left(\frac{G - \kappa y}{\sigma \sqrt{y}} \right) + \sigma \sqrt{y} \phi \left(\frac{G - \kappa y}{\sigma \sqrt{y}} \right) \right\}, \quad (41)$$

²We express the dependence of the optimal size of radical innovation on the gap G rather than the history \mathcal{H} from expositional clarity.

where $G = \bar{B} - \hat{B}$. The first-order condition is

$$\kappa \left(1 + \delta - \delta \Phi \left(\frac{G - \kappa y}{\sigma \sqrt{y}} \right) \right) + \frac{\delta \sigma}{2\sqrt{y}} \phi \left(\frac{G - \kappa y}{\sigma \sqrt{y}} \right) = c'(y). \quad (42)$$

The left-hand side of (42) approaches κ as G increases and, for fixed G , the left-hand side converges to $\kappa(1 + \delta)$ as $y \rightarrow +\infty$, and to κ as $y \rightarrow 0$. Also, the left-hand side is bounded above by

$$\kappa(1 + \delta) + \frac{\delta \sigma^2 \phi(1)}{2G},$$

which converges to $\kappa(1 + \delta)$, as G increases. Thus, $\lim_{y \rightarrow +\infty} c'(y) > \kappa(1 + \delta)$ (and $c''(0) > 0$, if $\kappa = c'(0)$) implies that there exists $\tilde{G} > 0$ such that $G > \tilde{G}$ implies $y^R(G) = 0$.

Differentiating (42) with respect to G yields

$$-\frac{\delta \phi}{2\sigma \sqrt{y}} \left[\kappa + \frac{G}{y} \right] < 0,$$

for all $G, y > 0$. This implies that $y^R(0) \geq y^R(G)$ for all $G > 0$. Repeating the argument used to prove Theorem 1, we conclude that radical experimentation ends in finite time a.s.

Suppose now that $\kappa > c'(0)$. An old agent experiments radically with a size equal to $y^{O,R} > 0$ if and only if $G \leq \xi$, as shown in the main text. We assume without loss of generality that $G > \xi$, so that an old agent does not experiment today and then the expected utility today of any $x > \hat{x}$ for a young agent is simply

$$\begin{aligned} & E[B(x) - c(x - \hat{x})] + \delta \{ E[B(x) + \xi | B(x) \geq \bar{B} - \xi] \text{Prob}(B(x) \geq \bar{B} - \xi) + \bar{B} \text{Prob}(B(x) < \bar{B} - \xi) \} \\ &= (1 + \delta)(\hat{B} + \kappa y) - c(y) + \delta \left\{ \sigma \sqrt{y} \phi \left(\frac{G - \xi - \kappa y}{\sigma \sqrt{y}} \right) + \xi + (G - \xi - \kappa y) \Phi \left(\frac{G - \xi - \kappa y}{\sigma \sqrt{y}} \right) \right\}. \end{aligned}$$

The first-order condition is

$$\kappa \left[1 + \delta - \delta \Phi \left(\frac{G - \xi - \kappa y}{\sigma \sqrt{y}} \right) \right] + \frac{\delta \sigma}{2\sqrt{y}} \phi \left(\frac{G - \xi - \kappa y}{\sigma \sqrt{y}} \right) = c'(y). \quad (43)$$

The right-hand side is always greater than or equal to κ . Since $G - \xi > 0$, the right-hand side converges to κ as $y \rightarrow 0$, and to $\kappa(1 + \delta)$ as $y \rightarrow +\infty$. Since also $c'(0) < \kappa$, there must exist an interior solution to the first-order equation (43). When $\lim_{y \rightarrow +\infty} c'(y) > \kappa(1 + \delta)$, the solution is unique for high values of G because the left-hand side of (43) approaches κ pointwise as G becomes arbitrarily large. As G increases, $y^R(G)$ approaches $y^{O,R}$: the optimal size of experimentation for a young agent converges to the optimal size for an old agent. This follows

from the fact that the first-order condition becomes, dropping negligible terms, $\kappa \approx c'(y)$. The maximized expected utility is approximately equal to

$$\begin{aligned} & (1 + \delta)(\hat{B} + \xi) + \delta \left[\kappa y^{O,R} \left(1 - \Phi \left(\frac{G - \xi - \kappa y^{O,R}}{\sigma \sqrt{y^{O,R}}} \right) \right) \right. \\ & \left. + \sigma \sqrt{y^{O,R}} \phi \left(\frac{G - \xi - \kappa y^{O,R}}{\sigma \sqrt{y^{O,R}}} \right) + (G - \xi) \Phi \left(\frac{G - \xi - \kappa y^{O,R}}{\sigma \sqrt{y^{O,R}}} \right) \right] \\ & \approx \hat{B} + \xi + \delta \bar{B}. \end{aligned}$$

Since we assumed that $G > \xi$, it follows that $\hat{B} + \xi + \delta \bar{B} < (1 + \delta)\bar{B}$. The only candidate for radical experimentation gives an expected payoff which is lower than what the young agent could get by simply exploiting. Thus, there exists $\tilde{G} > 0$ such that the young agent prefers exploitation for any gap greater than \tilde{G} .

In order to replicate the steps used to prove Theorem 1, we still need to show that there is an upper bound on the equilibrium size of experimentation. When $G > \xi$, the right-hand side of (43) is strictly decreasing in G . Thus, the unique positive solution of the first-order condition when $G = \xi$ provides the desired upper bound. When $G \leq \xi$, the old agent is experimenting with a fixed size $y^{O,R}$ (independent of the gap). Replicating the argument for $\kappa = 0$, one may show that, on the range $G \in [0, \xi]$, the value and size of radical experimentation are decreasing in G , providing an upper bound on the size of radical experimentation for a young agent, uniform over all histories. ■