Moonshots, Investment Booms, and Selection Bias in the Transmission of Cultural Traits

David Hirshleifer[1], Joshua B. Plotkin[2]

[1]Merage School of Business, UC Irvine, Irvine, CA 92617, USA

[2]Department of Biology, University of Pennsylvania, Philadelphia, PA 19104, USA

August 18, 2020

Abstract

Biased information about the payoffs received by others can drive innovation, risk-taking, and investment booms. We study this cultural phenomenon using a model based on two premises. The first premise is a tendency for large successes, and the actions that lead to them, to be more salient to onlookers than small successes or failures. The second premise is selection neglect – the failure of observers to adjust for biased observation. In our model, each firm in sequence chooses to adopt or to reject a project that has two possible payoffs, one positive and one negative. The probability of success is higher in the high state of the world than in the low state. Each firm observes the payoffs received by past adopters before making its decision, but there is a chance that an adopter's outcome will be censored, especially if the payoff was negative. Failure to account for biased censoring causes firms to become overly optimistic, leading to irrational booms in adoption. Booms may eventually collapse, or they may last forever. We describe these effects as a form of cultural evolution with adoption or rejection viewed as traits transmitted between firms. Evolution here is driven not only by differential copying of successful traits, but also by cognitive reasoning about which traits are more likely to succeed – quantified using the Price Equation to decompose the effects of mutation pressure and evolutionary selection. This account provides a new explanation for investment booms, merger and IPO waves, and waves of technological innovation.

1 Introduction

We study how biases in social transmission of information about the actions and payoffs of others induces innovation and risk-taking by firms. Our model is based on two premises. The first is the tendency for large successes, and the actions that led to them, to be more visible and salient to others than failures. For example, if 999 out of 1000 small start-ups fail completely, and one grows to become as large and successful as Google, each of the failures, being small, is seldom noticed and remembered, whereas discussion of the huge success becomes ubiquitous.

The second premise is the psychological phenomenon of *selection neglect*, the failure of observers to adjust for bias in the process that generates the data they observe.[1] We apply these two premises to business initiatives, such as creating a startup firm, undertaking a corporate 'moonshot' or 'sure bet' investment project, or making a large acquisition. We show that the dynamics of this evolutionary process can lead to boom and bust patterns, consistent with merger and IPO waves, and sudden waves of innovative technological activity.

[1]See, e.g., (Nisbett and Ross 1980), (Brenner, Koehler, and Tversky 1996), and, in the financial context, (Koehler and Mercer 2009).

We model a setting in which firms in sequence decide whether to adopt or to reject undertaking a project based on observation of the payoffs of previous firms. We consider *biased censorship*—a higher probability of observing successful projects and their outcomes than failed ones. Owing to selection neglect, a manager will have a biased assessment of the prospects of the project. We analyze how these factors shape boom/bust dynamics and long-term survival of different behaviors: adopting versus rejecting projects with different characteristics. We consider our model in the context of cultural evolution of financial traits in a population of firms.

In our setting, each firm in a sequence chooses whether to adopt or reject a project with two possible payoffs—a large success payoff or a negative failure payoff. The probability of success is higher in the high state of the world than in the low state, such that the project is profitable in expectation only in the high state. Each firm observes payoff outcomes of predecessors, but there is some probability that any given adopter's payoff is censored. If observers were rational, then firms would adopt in the long run only if the state of the world is high. However, owing to selection neglect, later firms tend to become overly optimistic about the state of the world, and so there can be a positive probability that all firms will adopt in the low state. In this state biased censorship can also exacerbate boom/bust patterns in which a long string of adopts occurs before ultimate collapse.

Viewing adopting or rejecting as a cultural trait which is transmitted between firms, we employ the Price Equation to decompose the evolution of adoption versus rejection of risky strategies into a selection component and a mutation component. Surprisingly, despite the central role of selection bias in the evolution of project choice in the model, under some action histories there is only mutation pressure, without evolutionary selection. This differs from cultural evolutionary models with direct copying, in which there is only selection, and it highlights the role of cognitive reasoning as a cultural evolutionary force.

We are not the first to examine selection bias and learning by firms. Denrell (Denrell 2003) also examines a setting in which observers neglect selection bias in the information they observe about firms. As in our model, failure is less likely to be observed than success. Denrell's focus is on how this biases learning about the traits that are characteristic of the upper tail of successful firms, which he argues will disproportionately consist of variance-increasing strategies. So Denrell concludes that selection bias will cause the spread of risky and unreliable management practices (Denrell 2003).

Our study differs in several ways. First, we focus on beliefs about the benefits of project adoption rather than about general managerial practices. Second, we allow for sequential choices and selection bias. In other words, we model how neglect of selection bias affects an arbitrary observer's *behavior*, not just beliefs; and how the observer in turn becomes the target of observation for the next agent, and so forth. This allows us to study the implications of selection bias for investment booms and collapses. Third, our modeling allows us to analyze the evolutionary process by which behavioral traits are transmitted across agents, in terms of selection and mutation pressure. Fourth, Denrell's focus is on how differences in variance biases choices, whereas we analyze payoff asymmetry, and show that greater moonshotness promotes project overadoption.

Han *et al.* (Han, Hirshleifer, and Walden 2020) examine a setting in which stock market investors randomly meet to discuss their strategies, and the probability that an investor reports the investor's return performance is increasing in return. An investor has an exogenous probability of copying another investor's strategy that is increasing with reported return, if the investor receives a report. As a result, high variance investing strategies spread through the population. Our focus here is on project choices by firms that update beliefs in a quasi-Bayesian fashion, based on sequential observation of a history of past payoffs by other firms. In contrast, in (Han, Hirshleifer, and Walden 2020) investors are randomly drawn to meet, and message receivers have an

Table 1: Payoff Probabilities

	Payoff Outcome (v)	
State (θ)	V	-1
H	p	$1-p$
L	$1-p$	p

exogenous switching probability based upon the single observation during that meeting. Our study also differs in deriving boom/bust dynamics, novel comparative statics about moonshotness, and decomposition of outcomes into the effects of selection versus mutation pressure.

2 The basic setting

Consider a setting in which agents in sequence choose between two actions, adopt or reject, and each learns about the state of the world from the payoff experiences of predecessors (see (Schlag 1998), (Bolton and Harris 1999), and (Cao, Han, and Hirshleifer 2011)). Reject generates a payoff of zero, whereas adopting generates binary possible payoffs. Let the possible actions for the n^{th} agent in the sequence, denoted I_n, be $a_n = A, R$ (adopt or reject a project). When an agent rejects, the payoff is zero. The binary state of the world is $\theta = H, L$, with prior probability q of state H. In each state, there are two possible net payoffs to adoption, $v = V > 0$ or $v = -1$. Project success has probability $p > 1/2$ in state H and $1-p$ in state L. Therefore high payoff is a symmetric binary signal about state, with payoff probabilities shown in Table 1.

We assume that when indifferent an agent rejects, and we impose parameter constraints such that adopting is, in expectation, strictly profitable in state H and unprofitable in state L. We let the expected profit from adopting conditional on state be Π^θ,

$$\begin{aligned} \Pi^H &\equiv E[v|\theta = H] = pV - (1-p) > 0 \\ \Pi^L &\equiv E[v|\theta = L] = (1-p)V - p < 0, \end{aligned} \tag{1}$$

which together imply that $p > \frac{1}{V+1}$ and $p > \frac{V}{V+1}$ If $V < 1$, only the first constraint is binding, and if $V > 1$, only the second one is.

Agents maximize expected profits conditional on their information, which derives from observing past payoffs. To avoid the trivial outcome that all agents always reject, we assume that the prior expected value of action A is positive,

$$\begin{aligned} E[v] &= q[pV - (1-p)] + (1-q)[(1-p)V - p] \\ &= q(2p-1)(V+1) + (1-p)V - p > 0. \end{aligned} \tag{2}$$

This implies that the first agent, I_1, adopts.

3 Full observability of past payoffs

If at any point an agent I_i's belief about the state becomes sufficiently pessimistic, I_i rejects, and I_i's payoff of 0 contains no information about state. Since I_{i+1} has no additional information, I_{i+1} also rejects, as do all later agents. So in the stochastic process of actions and payoffs, the action R is an absorbing outcome. With full observability of past payoffs, there is a random walk on beliefs, as measured by the log-likelihood ratio (LLR) between the two states, with a lower absorbing barrier (see e.g., (Cao, Han, and Hirshleifer 2011)).

Specifically, conditional on state, so long as agents adopt, the belief follows a random walk with drift. To see this, let d_i be the difference between the number of high payoff realizations (V) and the number of low realizations (-1) observed by agent I_i (i.e., through the payoff of I_{i-1}), where by convention, $d_1 = 0$. Then by Bayes' rule, since v realizations are conditionally independent given state $\theta = H$ or L,

$$\frac{P(H|d_i, v_i)}{P(L|d_i, v_i)} = \frac{P(v_i, d_i|H)}{P(v_i, d_i|L)}\frac{P(H)}{P(L)} = \frac{P(v|H)}{P(v|L)}\frac{P(H|d_i)}{P(L|d_i)}. \tag{3}$$

Taking logs,

$$\lambda_{i+1} \equiv \log\left(\frac{P(H|d_i, v_{i+1})}{P(L|d_i, v_i)}\right) = \log\left(\frac{P(v|H)}{P(v|L)}\right) + \log\left(\frac{P(H|d_i)}{P(L|d_i)}\right),$$

or

$$\lambda_{i+1} = \log\left(\frac{P(v|H)}{P(v|L)}\right) + \lambda_i. \tag{4}$$

So the LLR changes at each step adding the update increment $\log(\frac{P(v|H)}{P(v|L)})$, which is positive when $v = V$ and negative when $v = -1$, and which has the same absolute value in both cases. In state $\theta = H$, the probabilities of these increments are p for $v = V$ and $1 - p$ for $v = -1$. In state L, the probabilities are $1 - p$ for $v = V$ and p for $v = -1$.

The LLR (belief state) increases by $\log(p/(1 - p)) > 0$ or $\log((1 - p)/p) < 0$ at each step, by Eq. 4. It follows by simple induction that the LLR conditional upon all payoff observations is

$$\lambda_i \equiv \log\left(\frac{P(H|d_i)}{P(L|d_i)}\right) = d_i \log\left(\frac{p}{1-p}\right) + \log\left(\frac{q}{1-q}\right). \tag{5}$$

So the relevant payoff history from past adopts is fully summarized by d_i. (Rejects generate a deterministic payoff, and therefore are uninformative.)

The lower absorbing barrier is the cutoff on the LLR above which adoption has a positive expected payoff. To calculate this, observe that I_i adopts iff

$$E[v|d_i] = P(H|d_i)[pV - (1 - p)] + (1 - P(H|d_i))[(1 - p)V - p] > 0. \tag{6}$$

Bearing in mind Eq. 1, I_i adopts iff

$$\lambda_i = \log\left(\frac{P(H|d_i)}{1 - P(H|d_i)}\right) > \log\left(\frac{p - (1-p)V}{pV - (1-p)}\right) > 0. \tag{7}$$

Substituting Eq. 5 into Eq. 7, agent I_i adopts iff

$$d_i \log\left(\frac{p}{1-p}\right) + \log\left(\frac{q}{1-q}\right) > \log\left(\frac{p - (1-p)V}{pV - (1-p)}\right), \tag{8}$$

so the indifference cutoff on d is

$$d^* \equiv \frac{\log\left(\frac{p-(1-p)V}{pV-(1-p)}\right) - \log\left(\frac{q}{1-q}\right)}{\log\left(\frac{p}{1-p}\right)}. \tag{9}$$

Since an indifferent agent rejects, the first reject occurs when d_i reaches or crosses d^*, $d_i \leq \lfloor d^* \rfloor$. For example, if $d^* = -1.1$, the first reject occurs when $d_i = -2$.

The numerator of Eq. 9 is negative (see SI for details), which implies that $d^* < 0$. Intuitively, under the prior belief, an agent adopts, and so bad news ($d_i < 0$) is required for an agent to reject. In summary, the count of past high minus low payoffs follows a random walk with a lower absorbing barrier $\lfloor d^* \rfloor < 0$.

In general, for a random walk with a lower absorbing barrier at 0 and an up move probability $r \leq 1/2$ the probability of ultimate ruin starting from position d_0 is 1, and with $r > 1/2$, is $\left(\frac{1-r}{r}\right)^{d_0}$, which is decreasing in d_0 (Feller 1950). The standard walk as an equivalent transformation of our model in which the absorbing barrier $\lfloor d^* \rfloor < 0$ is increased to zero, and the starting position is accordingly increased to $d_0 = -\lfloor d^* \rfloor$). Here the probability of an up move corresponds to $p > 1/2$ or $1 - p$ in the H or L state respectively. We therefore have

Proposition 1 *With observation of all past actions and payoffs we have the following results:*

1. *In state L:*
   (a) *With probability 1 there exists an i such that for all $j \geq i$, agent I_j rejects.*
   (b) *The probability that agent I_n rejects is $(\mu^{-\lfloor d^* \rfloor} - \mu^{n-1})/(1 - \mu^{n-1})$, for $n > -\lfloor d^* \rfloor + 1$, where d^* is given by Eq. 9 and $\mu \equiv \frac{p}{1-p}$.*
2. *In state H:*
   (a) *The probability that all agents adopt is $1 - \nu^{-\lfloor d^* \rfloor}$, which is strictly between 0 and 1, where $\nu \equiv \frac{1-p}{p}$.*
   (b) *If not all agents adopt, then there exists an i such that for all $j \geq i$, agent I_j rejects.*
   (c) *The probability that agent I_n rejects is $(\nu^{-\lfloor d^* \rfloor} - \nu^{n-1})/(1 - \nu^{n-1})$ for $n > -\lfloor d^* \rfloor + 1$.*

Proof: Part 1a: In state L the LLR and d follow random walks with negative drift. It follows from standard properties of such walks that the lower absorbing barrier is reached with probability one. Part 1b: This follows from a standard result for random walks through step $n - 1$, bearing in mind that agent I_n only observes payoffs through agent I_{n-1}. Part 2a: In state H the random walk on d has positive drift. The first agent, I_1, adopts. A standard property of such a random walk is that there are strictly positive probabilities that such a walk ever or never hits a lower barrier. Part 2b: if the lower barrier is hit for some agent I_j, no further payoff information is generated, so all later agents also reject. Part 2c: This follows from a standard property of walks through step $n - 1$, bearing in mind that agent I_n only observes payoffs through agent I_{n-1}. ‖

The results of this Proposition are intuitive. For Part 1, if an arbitrarily large number of agents were to adopt, beliefs would converge to virtual certainty in state L, causing rejection. For Part 2, either early bad news makes adopt seem so bad that the reject barrier is crossed (in which case no further information is generated, so all subsequent agents reject), or else there is an infinite number of adopts, so that in the long run agents become highly confident of the H state.

4 Biased censorship and rational updating

Suppose now that high-payoff outcomes are more likely to be observed than low-payoff outcomes. Specifically, we assume that a payoff of $v = V$ is observed with probability $\delta \leq 1$, whereas a payoff of $v = -1$ is observed only with probability $\pi \leq \delta$. So the upside censorship probability is $1 - \delta \geq 0$ and the downside censorship probability is $1 - \pi \geq 0$. We assume strict inequalities, where the case of $\delta = \pi = 1$ is a benchmark for comparison.

We call greater censorship of downside outcomes *upside salience*, defined algebraically as the ratio $\beta \equiv \delta/\pi$. Intuitively, projects that succeed are associated with high scale of continuing economic transactions, which garners attention, whereas projects that fail tend to vanish.

We can think of censorship of observation of a firm's project outcome as meaning that the outcome is not reported conspicuously in the media. So we assume that censorship of an agent is universal: the agent's payoff is either visible to all successors or to none of them.[2]

Let $v_j = V$ or -1 be the realized payoff outcome of agent I_j if j adopts. Let O_{ij} denote the event that I_i observes I_j, where $i > j$, and $\overline{O}_{ij}$ denote the event that I_i does not observe I_j. Since I_1 adopts, observation by rational I_2 of I_1's payoff updates I_2's belief about state to (see SI for details):

$$
\begin{aligned}
P(\theta = H | O_{21}, V_1 = V) &= \frac{P(\theta = H, V_1 = V | O_{21})}{P(V_1 = V | O_{21})} = p \\
P(\theta = H | O_{21}, V_1 = -1) &= \frac{P(\theta = H, V_1 = -1 | O_{21})}{P(V_1 = -1 | O_{21})} = 1 - p.
\end{aligned}
$$

These expressions are independent of π and δ, because conditional on observation of a payoff, the greater censorship of low payoffs makes no difference for inferences. The only information contained in the sheer fact of observation is that the payoff is less likely to be -1. But conditional on a payoff that is directly observed, this information is redundant.

In contrast, the *absence* of an observation of an earlier agent contains useful information for agent I_i, as in the story in which detective Sherlock Holmes draws a key inference from "the dog that did not bark." For example, if I_2 does not observe I_1, then the rational inference is tilted toward state $\theta = L$, since the occurrence of no observation tends to comes from the payoff -1 (since $\pi < \delta$). The rational inference from no observation is (see SI for details)

$$
P(\theta = H | \overline{O}_{21}) = \frac{P(\theta = H | \overline{O}_{21})}{P(\overline{O}_{21})} = \frac{p(1-\delta) + (1-p)(1-\pi)}{2 - \delta - \pi}.
$$

5 Neglect of payoff-biased censorship

What if agents neglect the fact that there is biased censorship of their observations? We first consider how neglect of censorship affects short-term and long-term adoption dynamics. In Section 5 we examine basic comparative statics on long-run adoption. Finally, in Section 6, we examine the sources of cultural evolution toward adopt or reject using the Price Equation.

Neglect of biased censorship and adoption dynamics

When agents are unaware of censorship, if agent I_j observes $j' < j - 1$ past actions and payoffs, agent I_j mistakenly believes that he is the $(j'+1)^{\text{th}}$ agent in the decision queue. Intuitively, we expect that disproportionate censorship of low payoffs will make agents overly optimistic about project adoption.

We will show that the beliefs of an appropriate sequence of agents follows a random walk with drift, with an absorbing barrier that induces rejection. Consider the *subsequence* of the belief sequence that removes all agents whose payoffs are censored. This *uncensored subsequence* contains the beliefs of all agents whose payoffs matter for later agents. We will see that the beliefs (LLRs) of agents in this subsequence follow a random walk.

[2]As a result agents cannot glean any private information from the action choices of predecessors. This point is not relevant for our main analysis, since, in our setting with imperfect rationality, agents do not even realize that there is any such indirect information to be extracted.

Let a B (for Biased) superscript denote an imperfectly rational expectation. A biased agent mistakenly fails to condition on observation versus non-observation of payoffs when forming expectations. So for I_2 (see SI for details),

$$\begin{aligned} P^B(\theta = H|O_{21}, v_1 = V) &= \frac{P^B(\theta = H, V_1 = V)}{P^B(v_1 = V)} = p \\ P^B(\theta = H|O_{21}, v_1 = -1) &= \frac{P^B(\theta = H, v_1 = -1)}{P^B(v_1 = -1)} = 1 - p. \end{aligned}$$

We see that the beliefs *conditional on seeing V or -1* are the same as for a rational agent, independent of noncensorship probabilities δ and π. So neglecting the information implicit in the fact that a payoff *was* observed does not affect conditional beliefs.

In contrast, neglect of the information implicit in the absence of an observation crucially affects action dynamics and the evolution of beliefs. Unlike a rational agent, an inattentive agent does not draw inferences from censored observations.

When agents naively neglect non-observation, they update based on the observed difference between the numbers of high and low payoffs, just as in the case with no censorship. Specifically, let j be the index for the subsequence, i.e., a count in order of the uncensored agents. Let d_j be the difference in the number of V and -1 outcomes through the $(j-1)^{th}$ agent in this subsequence. Then Eq. 5 holds with i replaced with j.

However, neglect of censorship modifies the probabilities of an up or down move in beliefs conditional on state in the LLR walk. In the uncensored subsequence, the probabilities of an up move and down move in the next step are constant over time. So in the random walk for the uncensored subsequence, conditional on state $\theta = H$ the probabilities of an up move, p^{H*}, and of a down move, $1 - p^{H*}$, are (see SI for details)

$$\begin{aligned} p^{H*} &= P(v_1 = V|H, O_{21}) = \frac{p}{p + (1-p)/\beta} > p \\ 1 - p^{H*} &= \frac{1-p}{1 + p(\beta - 1)} < 1 - p, \end{aligned} \tag{10}$$

where the inequalities follow because upside salience $\beta > 1$. Note that the probability of an up move in the subsequence random walk conditional upon state H is increased by upside salience.

Similarly, in the subsequence random walk, conditional on state $\theta = L$ the probabilities of an up move and a down move respectively are

$$\begin{aligned} 1 - p^{L*} &= \frac{1-p}{1 - p + p/\beta} > 1 - p \\ p^{L*} &= \frac{p}{p + \beta(1-p)} < p, \end{aligned} \tag{11}$$

since $\beta > 1$.

There are two key differences in the evolution of beliefs from the case of no censorship. First, we must characterize beliefs for censored agents as well. These are determined trivially from the beliefs in the uncensored subsequence. Any censored agent's belief is identical to the belief of the next uncensored agent, since the next uncensored agent does not see intervening censored payoffs. Second, the probability of an up move in the LLR walk is higher with neglect of biased censorship. This effect can be arbitrarily large. As β becomes large ($\pi \approx 0$), both $p^{H*}, 1 - p^{L*} \to 1$. So with neglect of biased censorship, beliefs can tend to march upward, even in the L state where complete information would inevitably lead to rational rejection.

We summarize the evolutionary dynamics of adoption under biased censorship in the following Proposition.

Proposition 2 *Under biased censorship we have*

1. *In state L:*
   (a) *(Weak upside salience) If $\beta \leq p/(1-p)$, then there exists i such that for all $j > i$ agent I_j rejects.*
   (b) *(Strong upside salience) If $\beta > p/(1-p)$, then: (i) The probability that all agents adopt is $1-\zeta^{-\lfloor d^* \rfloor}$, where d^* is given by Eq. 9 and $\zeta \equiv \frac{p}{\beta(1-p)}$. This probability is strictly between 0 and 1, and is decreasing in π. (ii) If not all agents adopt, there exists i such that for all $j \geq i$ agent I_j rejects. (iii) The probability that I_n rejects is $(\zeta^{-\lfloor d^* \rfloor} - \zeta^{n-1})/(1-\zeta^{n-1})$ for $n > -\lfloor d^* \rfloor + 1$.*

2. *In state H:*
   (a) *The probability that all agents adopt is $1 - \eta^{-\lfloor d^* \rfloor}$, where $\eta \equiv \frac{(1-p)}{\beta p}$, which is strictly between 0 and 1 and decreasing in π.*
   (b) *If not all agents adopt, then there exists an i such that for all $j \geq i$, agents I_j rejects.*
   (c) *The probability of a reject within the first n agents is $(\eta^{-\lfloor d^* \rfloor} - \eta^{n-1})/(1-\eta^{n-1})$ for $n > -\lfloor d^* \rfloor + 1$.*

Proof:
Part 1a: The proof is essentially the same as Proposition 1, Part 1. The random walk has an equal up or down move, where in the L state, owing to selection bias, the up move and down move have respective probabilities $1 - p^{L*}$ and p^{L*} as in Eq. 11. So the random walk conditional on L has positive drift when $1 - p^{L*} > \frac{1}{2}$, or $\beta > \frac{p}{1-p}$. If instead $\beta \leq p/(1-p)$, the drift is negative or zero, so that the absorbing state, reject, is reached with probability 1. Part 1b: When upside salience is sufficiently strong ($\beta > p/(1-p)$), the LLR follows a random walk with positive drift, so the conclusions follow by the same reasoning as in Proposition 1, Part 2. The expression for the probability of reject within the first n agents follows by a standard result for random walks, where the probability of an up move in the subsequence random walk is given by Eq. 11. Parts 2a and 2b: For the H state, since $\beta > 1$ and $p > 1/2$, the LLR follows a random walk with positive drift. So conclusions similar to those of Part 1b about the probability of adopting forever follow by similar reasoning, where the probability of an up move conditional on the H state in the subsequence random walk is given by Eq. 10. Part 2c: The result follows by a standard result for random walks, where again the probability of an up move is given by Eq. 10. ‖

In sharp contrast with Proposition 1, owing to neglect of censorship of low payoffs, there is a positive probability that all agents will adopt, even when the project has negative expected value (state L). Moreover, the chance all agents adopt is greater in state H than without censorship.

Let $\hat{\beta} \equiv p/(1-p)$ be the cutoff value on upside salience such that with positive probability agents adopt in the long run even in state L. Clearly $\hat{\beta}$ is increasing with p, the superiority in the success rate in the high versus low state. When project success is highly sensitive to state, greater relative selection bias on the upside versus the downside is needed to induce long-run adoption in state L.

Proposition 2 shows that the model generates boom patterns, with strings of adoption, even in state L. Owing to upside salience, such booms occur more often when agents neglect censorship. In that case, there are irrational extra booms. Booms may be followed by collapse, or they may be sustained permanently. Adoption continues until (if ever) the preponderance of uncensored evidence, in the form of low payoffs, favors reject.

In some realizations there is no boom (a single adopt followed by all reject). In others there is a boom period followed by a collapse (several or many adopts and then all reject). In other cases there is a persistent boom (all adopt). In the L state, the boom component of the boom/collapse pattern is irrational relative to full knowledge of state.

Since agents do not know the state, booms and busts do not in themselves indicate bad decision making. Without censorship, all decisions are rational, but such mistakes still occur with positive probability. However, when there is biased censorship and selection neglect, there are *irrational* booms and busts. This is illustrated in Figure 1, which depicts realizations belief LLR over time. Without censorship, after two Adopts, the Reject barrier is reached, and all subsequent agents reject. In contrast, under censorship there is a temporary boom, until the late sequence of negative payoffs brings about collapse. This must eventually happen in the L state if upside salience is sufficiently weak. The bottom graph makes a similar comparison under a set of realizations in which, under censorship, the boom takes longer to collapse. Indeed, the boom may continue forever even in state L, provided upside salience is sufficiently strong.

Under selection neglect, the probability of an early string of adopts is increased—there are adoption bubbles. Suppose, for example, that there is relatively little censorship. By Proposition 2, with probability 1, in the L state eventually all agents reject. So censorship causes extra booms that would not occur were agents rational. If censorship is not too severe, these booms later crash. If censorship is more severe, matters are even worse—there are permanent mistaken booms. The model therefore offers a new explanation for real investment booms and busts (Chirinko and Schaller 2001)), IPO waves and overoptimism (Ritter 1991; Rajan and Servaes 1997; Lowry and Schwert 2002), and value-reducing merger waves (Moeller, Schlingemann, and Stulz 2005; Bouwman, Fuller, and Nain 2009)). This explanation differs from some past explanations that require payoff externalities (DeMarzo, Kaniel, and Kremer 2007) or shifts in investor sentiment (Gilchrist, Himmelberg, and Huberman 2005). There are, of course, other possible explanations as well.

Since key parameters of a model, such as β, can never be estimated perfectly, it is important to know whether model implications are robust to parameter variations. The qualitative implications of Propositions 1 and 2 are robust to small variations in β, except for the measure-zero set of critical values described in the propositions.

Comparative statics of long-run adoption

How does the upside payoff, the censorship probability, and the prior likelihood of success affect the chance of eventual adoption? To address these questions, in either state H or L, we analyze the partial derivatives of the chance of persistent adoption with respect to model parameters.

The long-run log rejection probability in the two states (where for state L, we impose the condition that $\beta > p/(1-p)$ so that ruin is not assured) is

$$\log(P(\text{Ever Reject}|\theta)) = \begin{cases} -(\lfloor d^* \rfloor) \log\left(\frac{1-p}{\beta p}\right) & \text{if } \theta = H \\ -(\lfloor d^* \rfloor) \log\left(\frac{p}{\beta(1-p)}\right) & \text{if } \theta = L. \end{cases} \tag{12}$$

Since $\log(p/(1-p))$ and $\log((1-p)/p)$ have opposite sign, it is evident that all the parameters except perhaps p and β (which reflects δ and π) have opposite directional effects on the long-run reject probabilities.

We use a version of Eq. 12 that ignores the floor function as a continuous approximation for the chance long-run rejection, after substituting Eq. 9 for d^*. To derive comparative statics, let

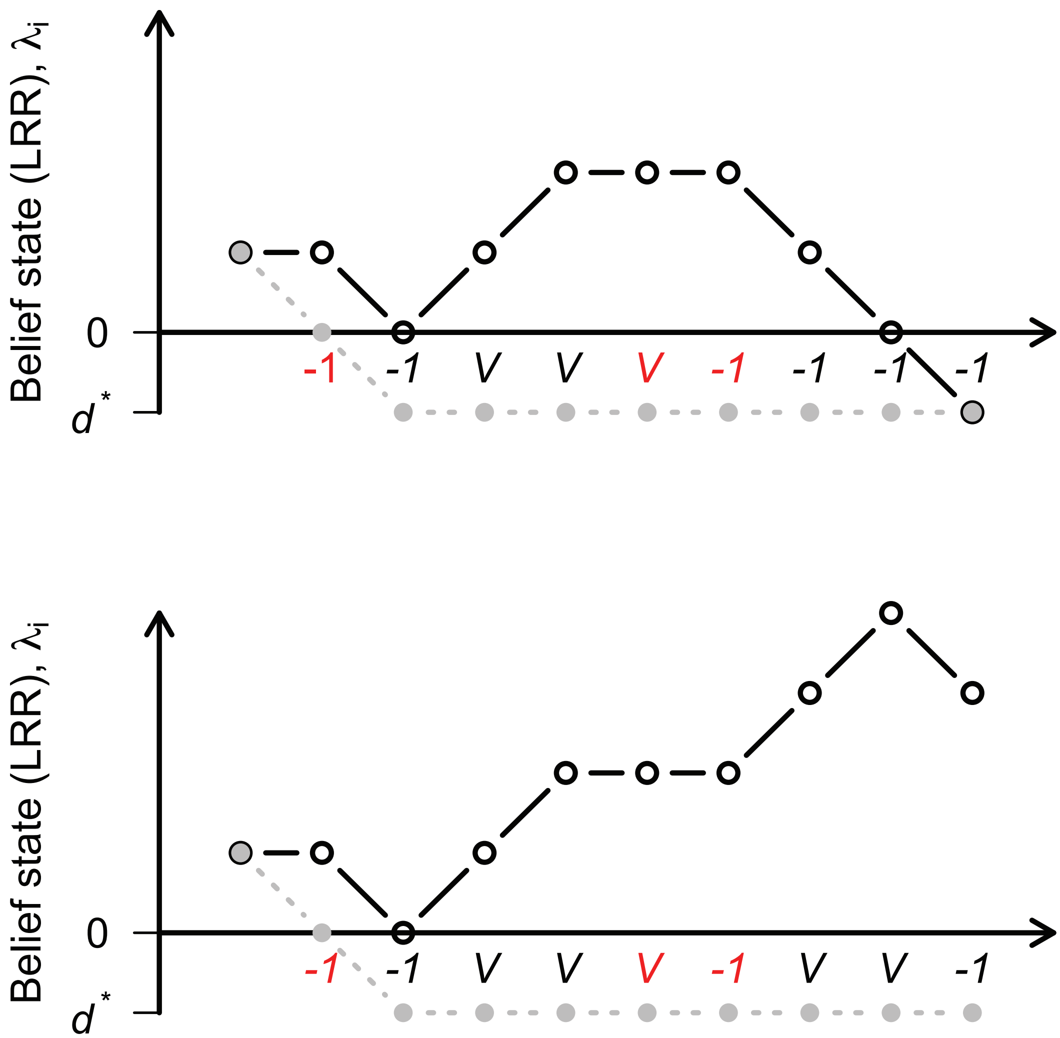

Figure 1: Booms of project adoption, with and without biased censorship. The top graph compares belief states of successive firms (x-axis) in the case of censorship (black) to the case of censorship (gray), under a set of payoff realizations in which the boom eventually collapses. The bottom graph makes a similar comparison under a set of realizations in which, under censorship, the boom takes longer to collapse, or may never collapse. Red payoffs correspond to censored outcomes.

R^H and R^L denote the log probabilities of ever rejecting in state H or L:

$$\begin{aligned} R^H &\equiv \log(P(\text{Ever Reject}|\theta = H)) \\ R^L &\equiv \log(P(\text{Ever Reject}|\theta = L)) \end{aligned} \tag{13}$$

We study the partial derivatives of R^H and R^L with respect to model parameters, p, q, V, and β (or π and δ separately) in regime with a positive chance of all agents adopting even in state L (i.e. $\beta > p/(1-p)$).

The effect of increasing upside salience follows immediately from differentiating Eq. 12 with respect to β. This gives $\frac{\partial R^H}{\partial \beta}, \frac{\partial R^L}{\partial \beta} < 0$. Since δ and π enter the reject probability expressions only through the definition of β, we also immediately have $\frac{\partial R^H}{\partial \delta}, \frac{\partial R^L}{\partial \delta} < 0$, and $\frac{\partial R^H}{\partial \pi}, \frac{\partial R^L}{\partial \pi} > 0$. So as upside salience becomes stronger (owing to either lower upside censorship or stronger downside censorship), the long-run chance of adoption under L increases. Intuitively, upside salience breeds overoptimism, reducing rejection probability.

High upside payoff, V, and high probability of the H state, q, both promote long-run adoption. These facts have opposite implications about the effects of moonshotness, which is by definition associated with both high V and low q. Some less obvious comparative statics, including those pertaining to interactive effects of parameters obtained from mixed partial derivatives, are derived and discussed in the SI. We summarize these results as follows.

Proposition 3 *In both states $\theta = H, L$, where for state L we require parameter values such that $R^L < 1$, we have the following comparative statics: $\frac{\partial R^\theta}{\partial \beta}, \frac{\partial R^\theta}{\partial \delta} < 0$ and $\frac{\partial R^\theta}{\partial \pi} < 0$; $\frac{\partial R^\theta}{\partial V}, \frac{\partial R^\theta}{\partial q} < 0$. Furthermore, in state L, for parameter values such that $R^L < 1$, we have: $\frac{\partial R^L}{\partial p} > 0$; $\frac{\partial^2 R^L}{\partial \beta \partial V}, \frac{\partial^2 R^L}{\partial \delta \partial V} < 0$ and $\frac{\partial^2 R^L}{\partial \pi \partial V} > 0$; $\frac{\partial^2 R^L}{\partial \beta \partial q} > 0$.*

Moonshots, sure bets, and upside salience

A moonshot is a project that has low probability of success, parameterized by the probability of the H state, q, and a high upside payoff V. This is the opposite of a 'sure bet' project, which has high q and low V. We expect upside salience β to be greater for moonshots than for sure bets. Both a low *ex ante* probability of success and a high conditional payoff makes success more surprising and newsworthy.

As stated in Proposition 3, long-run adoption is more likely when upside salience β is larger. So the model offers the empirical implication that moonshots will tend to be adopted more heavily than sure bet projects, even after controlling for net expected value.

We can study the effects of moonshotness formally by considering *excess* rejection probabilities in the two states. As before, we assume that $\beta > 1$, and, for the analysis of state L, focus on the case in which $\beta > p/(1-p)$ so that eventual rejection is not assured. Specifically, for two possible values of upside salience, β and $\overline{\beta}$ (where both satisfy the relevant inequality above), let the *excess log reject probability* relative to benchmark $\overline{\beta}$ in state H or L be defined as

$$\begin{aligned} e^H(\beta, \overline{\beta}) \equiv -\frac{R^{H,\beta}}{R^{H,\overline{\beta}}} &= \quad -\log\left(\frac{1-p}{\beta p}\right) / \log\left(\frac{1-p}{\overline{\beta} p}\right) \\ e^L(\beta, \overline{\beta}) \equiv -\frac{R^{L,\beta}}{R^{L,\overline{\beta}}} &= \; -\log\left(\frac{p}{\beta(1-p)}\right) / \log\left(\frac{p}{\overline{\beta}(1-p)}\right), \end{aligned} \tag{14}$$

where $R^{\theta,\beta}$ and $R^{\theta,\overline{\beta}}$ are the values of R^θ for two given values of β, and where the algebraic expressions follow from Eq. 12. The excess log reject probability captures the idea that, for any given values of q, V, and p, greater upside salience biases agents away from rejecting. So these expressions are decreasing with upside salience, β.[3] Notably, these are exact expressions, and they are independent of q and V.

To understand the effect of moonshotness in more depth, consider the probability of eventual rejection when varying V and q inversely while holding constant the prior expected payoff, $\overline{v} \equiv E(v)$. As argued above, an increase in moonshotness, captured here by an increase in V, increases upside salience β. Since the expressions in Eq. 14 have no direct dependence upon V and q, varying moonshotness affects these expressions *only* through its effect on upside salience β.

Specifically, differentiating Eq. 14 with respect to β yields

$$\begin{aligned} \frac{\partial e^H(\beta,\overline{\beta})}{\partial\beta} &= \frac{1}{\beta\log\left(\frac{1-p}{\overline{\beta}p}\right)} < 0 \\ \frac{\partial e^L(\beta,\overline{\beta})}{\partial\beta} &= \frac{1}{\beta\log\left(\frac{p}{\overline{\beta}(1-p)}\right)} < 0. \end{aligned} \tag{15}$$

So greater upside salience β decreases the excess rejection probability. By increasing upside salience, an increase in moonshotness decreases excess rejection, meaning it decreases rejection after controlling for the direct effects of V and q (other than their effect on β). Intuitively, moonshotness biases observation toward past successes rather than failures, which promotes adoption. In the SI, we discuss in greater detail the conceptual experiment underlying the comparative statics on excess log reject probabilities.

Salience and firm size

The noncensorship parameter π in the model is a proxy for the salience of downside outcomes, and δ for the salience of upside outcomes, where we have defined $\beta = \delta/\pi$ as the relative attention to the upside payoff (upside salience). Greater overall attention by media and observers will tend to increase both π and δ.

This effect will in general be asymmetric. For example, other things equal, large firms tend to receive much greater attention than small firms (O'Brien and Bushan 1990). In the limit, if a well-known firm such as Ford receives high attention to both its successes and failures, $\pi = \delta \approx 1$, so $\beta \approx 1$, i.e., there is no upside salience. In other words, for a large firm, failure of a major project can be notable enough to be reported upon in the media. In contrast, small start-ups often fail unnoticed. What is mainly reported in the media are extraordinary successes that start in garages and become tech giants. So other things equal, β will be higher for small firms than for large firms. The model therefore implies that the bias in favor of adopting risky projects will be especially strong for small start-ups. A further implication is that moonshots that are initiated with great fanfare and heavy investment may not generate as much mythology and overestimation as stories about firms that started in a garage.

If the effects of the model are stronger for small startups, then naive observers will strongly overestimate the probability of such startups succeeding, resulting in active and impetuous entrepreneurial activity. There is survey evidence that entrepreneurs are highly overoptimistic about their likely success (Cooper, Woo, and Dunkelberg 1988). Our model suggests that there will tend to be much less overoptimism about the innovative projects of large firms, and therefore less frequent undertaking of such projects.

[3] Since the benchmark denominator expression $R^{\theta,\overline{\beta}} < 0$, a negative sign is needed to ensure that higher $e^{\theta}(\beta,\overline{\beta})$ is indicative of higher reject probability.

6 Identifying sources of cultural evolution using the Price Equation

Our model is an evolutionary system in which later agents stochastically inherit the trait of adopt versus reject from earlier agents. If we designate some set of earlier agents as ancestral, and some set of later agents as descendant, then we can glean insight into the drivers of cultural evolution using the Price Equation (Price 1970). The Price Equation decomposes evolutionary change into selection and nonselection effects. The nonselection component is often called *mutation pressure*—the degree to which traits shift through the inheritance process instead of fitness-biased biased replication. The standard Price Equation applies to realizations; we take expectations to employ it in ex ante form.

The ex ante Price Equation

The insights provided by the Price Equation depend on how key categories are defined: ancestors, descendants, inheritance, and traits. We view *inheritance* in our setting as a potential causal dependence of the adopt/reject trait among descendants on the traits of ancestors. Under this interpretation there is no inheritance from a censored agent, because later agents derive no information from such an agent.

With this definition in mind, we now describe a stochastic version of the Price Equation. In the model, the history in general contains some mixture of high and low payoffs. Let the ancestral generation be denoted a and the descendant generation d. Let q^i be the frequency of type i in the a population, where $i = 0$ indicates reject and $i = 1$ indicates adopt. Let q'^i be the frequency of type i in the d population. We use overlines to denote expected frequencies, so let $\overline{q}'^i \equiv E[q'^i]$. In a realization where type $i = 0$ has zero frequency in a, fitness is undefined. We therefore use a form of the Price Equation emphasized by (Frank 1997) that contains ancestral and descendant frequencies instead of fitnesses.

The average trait values in the a and d populations, and the expected average trait in the d population conditional upon observed payoff information $\mathcal{F}$, are $\hat{z} \sum q^i z^i$, $\hat{z}' = \sum q'^i z'^i$, and $\overline{\hat{z}}'^{\mathcal{F}} \equiv E[\hat{z}'|\mathcal{F}] = \sum E[q'^i z'^i|\mathcal{F}]$. We study the change in trait value, $\Delta\hat{z} \equiv \hat{z}' - \hat{z}$ and the expected change, $\overline{\widehat{\Delta z}}^{\mathcal{F}} \equiv E[\hat{z}'|\mathcal{F}] - \hat{z}$.

Let $\Delta q^i \equiv q'^i - q^i$ be the frequency change due to natural selection, and let $\overline{\Delta q}^i \equiv \overline{q}'^{i\mathcal{F}} - q^i$ be the expected frequency change conditional upon information $\mathcal{F}$. Let the trait value change, and the expected trait value change conditional upon information $\mathcal{F}$ be $\Delta z^i \equiv z'^i - z^i$ and $\overline{\Delta z}^{i\mathcal{F}} \equiv \overline{z}'^{i\mathcal{F}} - z^i$.

The Price Equation decomposes average trait change into two terms

$$\Delta\hat{z} = \underbrace{\sum q'^i \left(\Delta z^i\right)}_{\text{mutation pressure}} + \underbrace{\sum \left(\Delta q^i\right) z^i}_{\text{selection}}, \tag{16}$$

as given by (Frank 1997). The expected change in the average population trait value is therefore

$$\overline{\widehat{\Delta z}}^{\mathcal{F}} = \underbrace{\sum \mathrm{E}[q'^i \left(z'^i - z^i\right) |\mathcal{F}]}_{\text{mutation pressure}} + \underbrace{\sum \mathrm{E}[(q'^i - q^i) z^i|\mathcal{F}]}_{\text{selection}}, \tag{17}$$

where in our context the sum is the 2 alleles, adopt and reject. We code these types with index $i = 0, 1$, which have ancestral alleles $z^0 = 0$ and $z^1 = 1$, where 0 indicates reject and 1 indicates adopt.

Consider, in state θ, a sequence of A or R realizations through agent I_n. We consider this population of n agents to be the ancestral generation, and we analyze the change or expected change in the average trait value to the descendant generation, I_{n+1}, which is a population of a single agent.

Under our definition of inheritance as a potential causal dependence of traits, an ancestral agent who is censored has no descendants. Furthermore, an agent with trait $i = 0$ rejects, resulting in a deterministic payoff of zero, which is uninformative to later agents. Since reject agents have no influence on later beliefs or actions, reject agents have no descendants. Finally, an *uncensored* agent with trait $i = 1$ does generate information that influences the information, and potentially the behavior, of the descendant I_{j+1}. So all uncensored agents collectively share the same, single descendant. That is, the descendant I_{j+1} inherits from *all* ancestral agents.[4] Since i refers to type in the above equations, the terms reflect the aggregated inheritance derived from all the agents of each type.

Selection and mutation pressure in project adoption

We study the expected evolution of traits conditional on the behaviors of early agents. We examine two conditionings: (i) all ancestral agents through I_n have adopted or (ii) not all ancestral agents adopted. Then we examine unconditional trait evolution in an illustrative special case.

Case 1: All ancestral agents adopted

Consider first the case in which, in state θ, all agents adopt through agent I_n. There is only type $i = 1$ in the a generation, so $q^0 = 0$ and $q^1 = 1$. Trivially, there can be no selection in the evolutionary sense, because the a population has only one allele.

Furthermore, there is stochastic mutation pressure wherein the $z^i = 1$ ancestors can map into a $z'^i = 0$ descendant. Observe that $q'^1 = 1$; all agents in d are descendants of type $i = 1$ in generation a.

Since it is not meaningful to study evolution when there are no ancestors, let $\mathcal{B}$ denote the event that there is at least one uncensored ancestor, which is a special case of the conditioning on general information set $\mathcal{F}$. We substitute into Eq. 17 to obtain

$$\begin{aligned} \overline{\widehat{\Delta z^{\mathcal{F}}}} &= \mathrm{E}[z'^1 - 1 | z_1 = \cdots = z_n = 1, \mathcal{B}] \\ &= P(z' = 1 | z_1 = \cdots = z_n = 1, \mathcal{B}) - 1, \end{aligned} \tag{18}$$

where the RHS expression comes solely from the mutation pressure term of the Price Equation. The superscript of $\mathcal{F}$ denotes conditioning on both $\mathcal{B}$ and on all past adopt.

What the Price Equation reveals here (Eq. 18) is that, even though there is *selection bias* on project payoffs that affects the evolution of the adopt/reject trait, there is no *selection* in the evolutionary sense. Selection bias induces cultural evolution by inducing mutation pressure in the trait value between ancestors and their descendants, rather than through selection on survival of ancestors.

This may seem counterintuitive, since a key driver of evolution here is that agents with low payoffs are "selected out." However, all ancestral agents have the same trait, $i = 1$. There is no variance for evolutionary selection to act upon. In contrast, censorship is based on whether the agents who adopted and experienced high versus low payoffs. That censorship decreases the tendency for I_{n+1} to switch to reject.

[4] This is consistent with versions of the Price Equation in which a descendant can have multiple ancestors. In the biological context, the number of ancestors is usually one or two for asexual or sexual reproduction, but in cultural contexts an agent can have many influencers.

The mutation pressure that drives expected trait change in Eq. 18 is negative, meaning there is always a chance that the descendant of an A agent will reject. Mutation pressure here is caused by cognitive reasoning, as later agents assess state based on the payoffs received by earlier agents. Notably, stronger selection bias *weakens* the pressure toward R. Moreover, the form of mutation pressure here differs from mutation pressure in genetic settings, in that its strength depends on the number of (uncensored) type-A agents in the ancestral generation. In genetic inheritance a mutation bias has a fixed strength and it applies to a single parent-offspring replication event. In our context, *all* the uncensored type-A ancestors are collectively the causal parents of a single offspring, and the greater their number, the greater the chance the offspring will convert to type R.

Case 2: A past reject has occurred

We now consider trait evolution conditional on at least one past reject. There is also at least one past adopt, since I_1 always adopts. Since the a generation has a mixture of adopt and reject alleles, there may be selection in the evolutionary sense.

Since there has been a past reject, all subsequent agents reject, so in the d generation agent I_{n+1} always rejects. In fact, with no additional complexity, we now generalize trivially so that the d generation includes any number of agents I_{n+1}, I_{n+2} and so forth. Also, since the outcome is deterministic, our stochastic version of the Price Equation reduces to the deterministic Price Equation as in Eq. 16.

Since a reject does not generate informative payoffs, and therefore has no causal effect on later agents, the reject type ($i = 0$) has no cultural descendants. In contrast, when an a agent adopts and is not censored, payoff information is transmitted to descendants and can influence behavior. Collectively, a sufficiently high number of low payoffs derived from uncensored past adopts *causes* the d agent(s) to reject. So as before, each agent in the d generation descends from *all* the agents of type $i = 1$, and only type $i = 1$ has descendants. So by definition, the descendant frequencies are $q'^0 = 0$, $q'^1 = 1$.

Since, under Case 2 conditioning, the descendants of type $z^1 = 1$ always reject, we have $z'^1 = 0$. In other words, trait $i = 1$ has *perfect negative heritability.* This is reflected in the mutation pressure component of the Price Equation. Substituting these values into the Price Equation gives

$$\overline{\widehat{\Delta z}}^{\mathcal{F}} = \underbrace{-1}_{\text{mutation pressure}} + \underbrace{1 - q^1}_{\text{selection}} \tag{19}$$

where $\mathcal{F}$ refers in this case to the conditioning on at least one past reject and at least one uncensored adopt. The initial mean value of z^i in a is $q^0(0) + (1 - q^0)1 = 1 - q^0$. The initial a mean and the change in mean are negatives of each others, which reflects the fact that all descendants reject. (mean trait value of zero).

The decomposition shows that evolution toward the reject allele derives from the opposition of two strong effects: evolutionary selection and mutation pressure. There is strong selection for A, since only A types leave descendants. However, this is overwhelmed by even stronger mutation pressure toward R. Specifically, evolutionary selection results in no descendants of the $i = 0$ type, yielding a positive selection term. But mutation pressure in the descendants of the $i = 1$ type is overwhelming; all of them shift from $z^1 = 1$ to $z'^1 = 0$. This generates the -1 term. So the R allele becomes fixed in the d generation.

The mutation pressure in this setting is more extreme than in previous case with no rejects in the ancestral generation (Eq 18), because this setting is deterministic: the descendent generation definitively rejects, and the rejection assuredly arose by mutation from an ancestor who had adopted. Furthermore, the strength of selection (i.e. change in mean phenotype due to differential survival) is strongest when q^1 is small – that is, when the ancestral generation has only a small frequency of adopters – because here there is a greater differential contribution of adopters versus rejectors to the descendent generation.

The censorship parameter *pi* does not enter Eq. 19. Intuitively, censorship affects who becomes an A or R in the a generation, but once we condition on an R being present in the a generation, this determines that R will be fixed in the d generation. So biased censorship does not affect the strength of evolutionary selection.

Case 3: Without conditioning on whether past rejects

In the SI, we also perform a Price Equation decomposition in a case with minimal conditioning—only on the presence of at least one uncensored ancestor. This allows for the cases of All Past Adopt (Case 1) or Some Past Reject (Case 2) as possible realizations. The insights from Cases 1 and 2 carry through to this minimal-conditioning case. There is still opposition between the effects of selection and mutation pressure, with selection favoring adopt and mutation pressure favoring reject.

Taken together, the three applications of the Price Equation above provide a notable contrast between cultural and genetic evolution. Moreover, evolution in our context also differs from trait dynamics in other behavioral settings, such as in evolutionary game theory where inheritance is determined by direct copying (Nowak 2006). Accurate copying leads only to selection, i.e., differential reproduction of traits. In contrast, agents here process information thoughtfully, which enriches the causality of trait transmission. Information transmission from parent to offspring results in mutation pressure, which can even overwhelm selection.

This application of the Price Equation offers three lessons. First is that care is needed in drawing interpretations about cultural transmission in terms of selection. It is tempting to view cultural transmission as simply a matter of differential survival of traits that are more effective at reproducing themselves. This form of selection can certainly occur, but cultural transmission can also take very different forms. As we have seen, one of these forms is a systematic, endogenous mutation pressure arising from cognition. Such pressure can operate even when there is no selection (Case 1), and it can overwhelm selection when the two forces are opposed (Case 2).

Second is that the Price Equation decomposition can apply to many social economic models, so long as some agents are influenced by other agents, and a causal linkage between actions is used to specify inheritance. The third lesson is that, unlike in evolutionary game theory (Nowak 2006), mutation pressure will often be important in these settings, because agents engage in cognitive reasoning about unknown states.

7 Discussion

Biased information about others can profoundly influence investment risk-taking, which we have studied as a cultural trait transmitted among firms. We have shown that when low-payoff outcomes are censored with higher probability than high-payoff outcomes, firms that do not account for this censorship bias become overly optimistic and undertake projects too often. This causes booms of overadoption, followed either by a eventual bust or by permanent long-run adoption even in a state where complete information would assuredly lead to (rational) rejection of a risky project.

These dynamics are a form of cultural evolution in which parentage reflects causality. This is a richer form of cultural evolution than simple copying of successful traits, because it allows for cognitive reasoning about which traits are more likely to yield high payoff based on (biased) observations of prior outcomes.

Even for given size of project, some types of firms (such as large firms) tend to receive more attention than others (e.g., greater coverage by financial and media analysts). Since attention probabilities are bounded above by one, the upside salience of high-profile firms is necessarily limited; even their failures are noticed.

Specifically, when attention is high, censorship of low outcomes is less likely, which reduces overadoption. For example, if a huge moon-shot initiative by a large established firm (e.g., self-driving cars) were to fail, this could be as conspicuous as a success. So the model implies that overadoption will be more pronounced among startups and small firms than among large established enterprises. The extreme difference in visibility of success versus failure among startups provides a new explanation for the overoptimistic expectations of entrepreneurs, and for the empirical anomaly of low returns to private equity (Moskowitz and Vissing-Jørgensen 2002).

Some firms receive higher attention for reasons other than sheer size. For example, as a creator of innovative consumer products, Apple has long been a magnet for attention, starting even before it became a giant. Firms in consumer product businesses tend to attract greater attention than infrastructure firms. A failure of a project by a firm in a high-attention sector may be more salient than failure in a sector that receives little attention.

When a firm's decision making is influenced by observations of other firms, adopting or rejecting a project is a cultural trait transmitted with bias between firms. We use the Price Equation to decompose such transmission into a selection component and a mutation component. Notably, even though misperceptions of managers are driven by selection bias in what they observe, the Price Equation reveals that behavioral evolution is driven by the opposing effects of mutation pressure and evolutionary selection, and that in some conditions there is only mutation pressure, without evolutionary selection. This contrasts sharply with a large domain of cultural evolutionary models with accurate copying, in which there is only selection.

The result of excessive adoption in our model is relative to a rational firm-level optimum. However, there are, in general, positive externalities to research and innovation. So excessive adoption that is unprofitable at the firm level, may be welfare-increasing at the social level. Innovative and moonshot projects may generate especially high externalities. On the other hand, there are also undesirable innovations such as patent trolling and the use of hijacked airplanes as weapons.

We assume *upside salience*: that the probability that a high payoff is observed by others be greater than the probability that a low payoff is observed, but less than one. We expect moonshots to have high upside-salience, because a rare, very high payoff is especially noticeable. In consequence, the model implies that overadoption will be more severe for moonshot projects. We also expect 'sexy' projects (project that are innovative, fun and exciting, such as self-driving cars) to have high upside-salience. For given upside cash flow, people especially like to hear about projects that they feel will change life in vivid ways. And so there may be a tendency to over-invest in such sexy projects.

We have shown that upside salience causes managers to overvalue moonshot projects, resulting in over-adoption and boom/bust patterns. For similar reasons, security market investors may overvalue "lottery stocks" (stocks with positive skewness), consistent with evidence from stock returns as summarized in (Han, Hirshleifer, and Walden 2020).

Our main focus has been investment at the firm level, but our approach can also be applied at larger scales to explain industry-level or aggregate-level investment boom/bust patterns. Industries may differ in upside salience of payoff outcomes, in part owing to differences in average firm size. So our model suggests that boom/bust patterns can be much more pronounced in some industries than in others. This is an interesting direction for empirical testing.

Stepping beyond the model somewhat, our approach also has organizational implications, including an advantage to managers of recruiting team members who are less heavily censored in observing others. Examples would include directors and venture capitalists who have broader direct experience, rather than observation through the media, than the manager or entrepreneur in past projects, especially for small startups.

ACKNOWLEDGMENTS. We thank participants at the Conference on Evolutionary Models of Financial Markets, MIT Laboratory for Financial Engineering (virtual), June 2020, and especially the discussant, Terry Burnham, for insightful comments.

References

Bolton, P. and C. Harris (1999, February). Strategic experimentation. *Econometrica 67*(2), 349–374.

Bouwman, C., K. Fuller, and A. Nain (2009, February). Market valuations and acquisition quality: Empirical evidence. *Review of Financial Studies 22*(2), 633–679.

Brenner, L. A., D. J. Koehler, and A. Tversky (1996, March). On the evaluation of one-sided evidence. *Journal of Behavioral Decision Making 9*(1), 59–70.

Cao, H. H., B. Han, and D. Hirshleifer (2011, July). Taking the Road Less Traveled by: Does Conversation Eradicate Pernicious Cascades? *Journal of Economic Theory 146*(4), 1418–1436.

Chirinko, R. S. and H. Schaller (2001, June). Business fixed investment and "bubbles": The Japanese case. *American Economic Review 91*(3), 663–680.

Cooper, A. C., C. W. Woo, and W. C. Dunkelberg (1988, Spring). Entrepreneurs' perceived chances for success. *Journal of Business Venturing 3*(2), 97–108.

DeMarzo, P., R. Kaniel, and I. Kremer (2007, September). Technological innovation and real investment booms and busts. *Journal of Financial Economics 85*(3), 735–754.

Denrell, J. (2003, May). Vicarious learning, undersampling of failure, and the myths of management. *Organization Science 14*(3), 227–243.

Feller, W. (1950). *An Introduction to Probability Theory and Its Applications.* New York, NY: Wiley.

Frank, S. A. (1997). The Price equation, Fisher's fundamental theorem, kin selection, and causal analysis. *Evolution 51*(6), 1712–1729.

Gilchrist, S., C. P. Himmelberg, and G. Huberman (2005, May). Do stock price bubbles influence corporate investment? *Journal of Monetary Economics 52*(4), 805–827.

Han, B., D. Hirshleifer, and J. Walden (2020, June). Social Transmission Bias and Investor Behavior. *Journal of Financial and Quantitative Analysis*. Forthcoming.

Koehler, J. J. and M. Mercer (2009, July). Selection neglect in mutual fund advertisements. *Management Science 55*(7), 1107–1121.

Lowry, M. and G. W. Schwert (2002, June). IPO market cycles: Bubbles or sequential learning? *Journal of Finance 57*, 1171–1200.

Moeller, S. B., F. P. Schlingemann, and R. M. Stulz (2005, April). Wealth destruction on a massive scale? A study of acquiring-firm returns in the recent merger wave. *Journal of Finance 60*(2), 757–782.

Moskowitz, T. J. and A. Vissing-Jørgensen (2002, September). The returns to entrepreneurial investment: A private equity premium puzzle? *American Economic Review 92*(4), 745–778.

Nisbett, R. and L. Ross (1980). *Human Inference: Strategies and Shortcomings of Social Judgment*. Englewood Cliffs, NJ: Prentice-Hall.

Nowak, M. A. (2006). *Evolutionary Dynamics: Exploring the Equations of Life*. Cambridge, Massachusetts: Belknap Press of Harvard University Press.

O'Brien, P. and R. Bushan (1990). Analyst following and institutional ownership. *Journal of Accounting Research 28*, 55–76.

Price, G. R. (1970, August). Selection and covariance. *Nature 227*, 520–521.

Rajan, R. G. and H. Servaes (1997, June). Analyst following of initial public offerings. *Journal of Finance 52*(2), 507–530.

Ritter, J. R. (1991, March). The long-run performance of initial public offerings. *Journal of Finance 46*(1), 3–27.

Schlag, K. H. (1998, January). Why imitate and if so how? A boundedly rational approach to multi-armed bandits. *Journal of Economic Theory 78*(1), 130–156.

Supplementary Information for

Moonshots, Investment Booms, and Selection Bias in the Transmission of Cultural Traits

David Hirshleifer, Joshua B. Plotkin

David Hirshleifer
E-mail: david.h@uci.edu

This PDF file includes:

Supplementary text
Table S1

Supporting Information Text

1. Updating under full observation: additional details

In the main text we claim that d^* is negative. The reasoning is as follows. Under the prior belief, adopting is profitable. Since $q(2p-1)(V+1) > 0$, we have $1 - q < \frac{pV-(1-p)}{(2p-1)(V+1)}$. The numerators and denominator in these two expressions are positive, so dividing gives $\log(\frac{p-(1-p)V}{pV-(1-p)}) - \log(\frac{q}{1-q}) < 0$, which is the numerator of our expression for d^*.

2. Updating under biased censorship: derivations

Rational updating under biased censorship: derivations. Let $v_j = V$ or -1 be the realized payoff outcome of agent I_j if j adopts. Let O_{ij} denote the event that I_i observes I_j, where $i > j$, and $\overline{O}_{ij}$ denote the event that I_i does not observe I_j. Since I_1 adopts, observation by rational I_2 of I_1's payoff updates I_2's belief about state to

$$\begin{aligned} P(\theta = H|O_{21}, V_1 = V) &= \frac{P(\theta = H, V_1 = V|O_{21})}{P(V_1 = V|O_{21})} = \frac{\left(\frac{1}{2}\right)p\delta}{\left(\frac{1}{2}\right)p\delta + \left(\frac{1}{2}\right)(1-p)\delta} = p \\ P(\theta = H|O_{21}, V_1 = -1) &= \frac{P(\theta = H, V_1 = -1|O_{21})}{P(V_1 = -1|O_{21})} = \frac{\left(\frac{1}{2}\right)(1-p)\pi}{\left(\frac{1}{2}\right)(1-p)\pi + \left(\frac{1}{2}\right)p\pi} \\ &= 1 - p. \end{aligned} \quad [1]$$

By contrast, if I_2 does not observe I_1, the rational inference from no observation is:

$$\begin{aligned} P(\theta = H|\overline{O}_{21}) &= \frac{P(\theta = H|\overline{O}_{21})}{P(\overline{O}_{21})} \\ &= \frac{P(H, v_1 = V)(1-\delta) + P(H, v_1 = -1)(1-\pi)}{P(H, v_1 = V)(1-\delta) + P(H, v_1 = -1)(1-\pi) + P(L, v_1 = V)(1-\delta) + P(L, v_1 = -1)(1-\pi)} \\ &= \frac{\left(\frac{1}{2}\right)p(1-\delta) + \left(\frac{1}{2}\right)(1-p)(1-\pi)}{\left(\frac{1}{2}\right)p(1-\delta) + \left(\frac{1}{2}\right)(1-p)(1-\pi) + \left(\frac{1}{2}\right)(1-p)(1-\delta) + \left(\frac{1}{2}\right)p(1-\pi)} \\ &= \frac{p(1-\delta) + (1-p)(1-\pi)}{2 - \delta - \pi}. \end{aligned} \quad [2]$$

Neglect of biased censorship and adoption dynamics: derivations. Let a B superscript denote an imperfectly rational (biased) expectation. A biased agent mistakenly drops the conditioning on observation versus non-observation of payoffs in forming expectations. So for I_2,

$$\begin{aligned} P^B(\theta = H|O_{21}, v_1 = V) &= \frac{P^B(\theta = H, V_1 = V)}{P^B(v_1 = V)} \\ &= \frac{P^B(\theta = H, V_1 = V)}{\left(\frac{1}{2}\right)P^B(v_1 = V|\theta = H) + \left(\frac{1}{2}\right)P^B(v_1 = V|\theta = L)} \\ &= \frac{\left(\frac{1}{2}\right)p}{\left(\frac{1}{2}\right)p + \left(\frac{1}{2}\right)(1-p)} = p \\ P^B(\theta = H|O_{21}, v_1 = -1) &= \frac{P^B(\theta = H, v_1 = -1)}{P^B(v_1 = -1)} \\ &= \frac{\left(\frac{1}{2}\right)(1-p)}{\left(\frac{1}{2}\right)(1-p) + \left(\frac{1}{2}\right)p} = 1 - p. \end{aligned} \quad [3]$$

In the random walk for the uncensored subsequence, conditional on state $\theta = H$ the probabilities of an up move, p^{H*}, and of a down move, $1 - p^{H*}$, are

$$\begin{aligned} p^{H*} &= P(v_1 = V|H, O_{21}) = \frac{P(v_1 = V, H, O_{21})}{P(v_1 = V, H, O_{21}) + P(v_1 = -1, H, O_{21})} \\ &= \frac{\left(\frac{1}{2}\right)P(v_1 = V, O_{21}|H)}{\left(\frac{1}{2}\right)P(v_1 = V, O_{21}|H) + \left(\frac{1}{2}\right)P(v_1 = -1, O_{21}|H)} \\ &= \frac{p\delta}{p\delta + (1-p)\pi} \\ &= \frac{p}{p + (1-p)/\beta} > p \\ 1 - p^{H*} &= \frac{1-p}{1 + p(\beta - 1)} < 1 - p, \end{aligned} \quad [4]$$

where the inequalities follow because upside salience $\beta = \delta/\pi > 1$.

3. Derivation of comparative statics

In the main text we derived the inequality

$$\log\left(\frac{p-(1-p)V}{pV-(1-p)}\right) - \log\left(\frac{q}{1-q}\right) < 0. \quad [5]$$

We can show several simple conclusions. For state H, $\frac{\partial R^H}{\partial V} < 0$ and $\frac{\partial R^H}{\partial q} < 0$. In other words, as the upside reward V of the project increases, or the prior likelihood of state H increases, the probability of long-run adoption increases. Specifically, differentiating the expression for R^H from the main text gives

$$\frac{\partial R^H}{\partial q} = \frac{\log\left(\frac{1-p}{\beta p}\right)}{q(q-1)\log(\frac{p}{1-p})}. \quad [6]$$

The denominator is positive, because $p > \frac{1}{2}$, and the numerator is negative, because $\beta > p/(1-p)$, so $\frac{\partial R^H}{\partial q} < 0$. Similarly,

$$\frac{\partial R^H}{\partial V} = \frac{(2p-1)\log\left(\frac{1-p}{\beta p}\right)}{(p(1+V)-V)(p(1+V)-1)\log\left(\frac{p}{1-p}\right)}. \quad [7]$$

The numerator is negative, again, whereas each term in the denominator is positive (because $p > \frac{V}{1+V}$ and $V > 1$ and $p > \frac{1}{2}$), and so $\frac{\partial R^H}{\partial V} < 0$.

Likewise, for state L, when $R^L < 1$, we have $\frac{\partial R^L}{\partial p} > 0$, $\frac{\partial R^L}{\partial V} < 0$, and $\frac{\partial R^L}{\partial q} < 0$. Specifically,

$$\frac{\partial R^L}{\partial q} = -\frac{\log\left(\frac{\beta(1-p)}{p}\right)}{(1-q)q\log\left(\frac{p}{1-p}\right)}. \quad [8]$$

The numerator is positive for $\beta > p/(1-p)$ and the denominator positive for $p > 1/2$, and so $\frac{\partial R^L}{\partial q} < 0$.

Likewise

$$\frac{\partial R^L}{\partial V} = \frac{(2p-1)\log\left(\frac{p}{\beta(1-p)}\right)}{[p(1+V)-V][]p(1+V)-1]\log\left(\frac{p}{1-p}\right)}. \quad [9]$$

In this case the denominator is again positive, and numerator again negative, under the condition $\beta > p/(1-p)$, so that $\frac{\partial R^L}{\partial V} < 0$.

Likewise,

$$\frac{\partial R^L}{\partial \beta} = \frac{\partial R^H}{\partial \beta} = \frac{\log\left(\frac{q}{1-q}\right) - \log\left(\frac{p(1+V)-V}{p(1+V)-1}\right)}{\frac{1}{\beta}\log\left(\frac{p}{(1-p)}\right)}. \quad [10]$$

Here the denominator is positive, because $p > 1/2$, and the second term in the numerator is also positive because $V > 1$. Although the sign of the numerator in general depends on q, the restrictions on q imposed by (2) assure that the numerator is positive by Eq. (5). So $\frac{\partial R^L}{\partial \beta} = \frac{\partial R^H}{\partial \beta} < 0$.

Likewise

$$\frac{\partial R^L}{\partial p} = \frac{\frac{\zeta\log\left(\frac{p}{\beta(1-p)}\right)}{(p-1)p} - \log\left(\frac{p}{1-p}\right)\left(\frac{(V^2-1)\log\left(\frac{p}{1-p}\right)}{[(p-1)V+p](pV+p-1)} + \frac{\zeta}{(p-1)p}\right)}{\log^2\left(\frac{p}{1-p}\right)} > 0 \quad [11]$$

where

$$\zeta \equiv \log\left(\frac{q}{1-q}\right) - \log\left(\frac{p(1+V)-V}{p(1+V)-1}\right) > 0 \quad [12]$$

by Eq. 5.In other words, higher p promotes rejection because payoffs become more accurate indicators of the actual state, L. Agents also understand that payoffs are more accurate, and therefore update more strongly to payoff outcomes. This increases the sizes of up- and down-moves in the LLR random walk, reducing the mean number of steps required to reach the absorbing barrier. This effect also promotes rejection in the long run.*

Some less obvious comparative statics are obtained by examining interactive effects of parameters by inspecting mixed partial derivatives. These provide more distinctive empirical implications of the model. When we vary both upside salience and upside payoff, we obtain

$$\frac{\partial^2 R^L}{\partial\beta\partial V} = -\frac{2p-1}{\beta[p(1+V)-V][p(1+V)-1]\log\left(\frac{p}{1-p}\right)} < 0. \quad [13]$$

* In contrast, in the H state the two effects are opposing, so the comparative statics on p is ambiguous.

Since $\beta = \delta/\pi$, and β and π enter into the expression for R^L only through β, it follows immediately that $\partial^2 R^L/(\partial\delta\partial V) < 0$ and $\partial^2 R^L/(\partial\pi\partial V) > 0$.

To understand this, first consider the effects of V and π on the chance of ever rejecting individually. Rejection occurs when the random walk d'_k (the net count of up versus down steps observed by agent I_k) hits the lower boundary $d^* < 0$. Increasing the upside payoff V decreases d^*, making rejection less likely ($\partial R^L/\partial V < 0$). Increasing π (reducing censorship) increases the chance of a downward step, making rejection more likely ($\partial R^L/\partial\pi > 0$). However, if V is already very large, the chance of ever reaching the lower boundary $d^* \ll 0$ is almost zero, and so the marginal effect of an increase in π is small. So the cross partial on R^L is positive. A similar intuition applies with opposite sign for the cross partials with respect to β or δ.

Likewise,

$$\frac{\partial^2 R^L}{\partial\beta\partial q} = \frac{1}{\beta q(1-q)\log\left(\frac{p}{1-p}\right)} > 0. \qquad [14]$$

The intuition here is essentially the same as that for the cross partial of β with V discussed above.[†]

4. Motivating conceptual experiment for comparative statics on excess log reject probabilities

The conceptual experiment underlying the comparative statics on the excess log reject probabilities is a differences-in-differences comparison. The starting point is a pair of projects with the same q and V values but different β values, β and $\overline{\beta}$. Taking the partial derivative with respect to β involves a comparison with another pair of projects with beta values of $\beta' > \beta$ and $\overline{\beta}$, where $\beta' - \beta$ approaches zero. This pair of projects has values q' and V', which are consistent with higher moonshotness. The result of this experiment therefore answers the question of how much a small increase in β that derives from a change in moonshotness changes the reject probability, after adjusting for the direct effect of the shift in q and V values from (q, V) to (q', V') ("direct" meaning other than through their effect on β). Specifically, since $e^{\theta}(\beta, \overline{\beta})$ is independent of the q and V values, this comparison accommodates the second pair of projects having any arbitrary different common set of q and V values, q' and V'.

This can be seen in more detail from Table S1, which shows the four projects involved. Specifically, this adjustment is reflected in the fact that when we compare a project with (β, q, V) to a project with (β', q', V'), the normalization benchmark factor in the denominator shifts at the same time from a project with $(\overline{\beta}, q, V)$ to a project with $(\overline{\beta}, q', V')$.

We apply this general experiment to the specific values of β', q' and V' implied by varying moonshotness. We have argued that β is increasing with moonshotness, so that if we increase V while decreasing q parametrically to hold expected payoff constant, β decreases. In other words, we can define the function $\beta_\kappa(V; p, \overline{v})$ for the relation between β and moonshotness as parameterized by V, and assume that $\partial\beta_\kappa(V)/\partial V > 0$. As discussed earlier, we also have the parametric relationship $\partial q_\kappa(V)/\partial V < 0$. So in the experiment above, $q' = q_\kappa(V + \Delta V; p, \overline{v})$, and $\beta' = \beta_\kappa(V + \Delta V; p, \overline{v})$.

5. Case 3 of Price Equation: No conditioning on whether there is a past reject

Here we perform a Price equation decomposition with minimal conditioning—only on the presence of at least one uncensored ancestor (event $\mathcal{B}$). This allows for the cases of All Past Adopt (Case 1) or Some Past Reject (Case 2) as possible realizations.

From the expectational Price equation, recognizing that $z^0 = z'^0 = 0$, $z^1 = 1$, so $\sum q^i z^i = q^1$, and $\sum \mathrm{E}[q'^i z^i|\mathcal{B}] = \mathrm{E}[q'^1 z^1|\mathcal{B}]$, we obtain

$$\begin{aligned} \overline{\widehat{\Delta z}}^{\mathcal{B}} &= \underbrace{\mathrm{E}[q'^1\left(z'^1 - z^1\right)|\mathcal{B}]}_{\text{mutation pressure}} + \underbrace{\mathrm{E}[(q'^1 - q^1)z^1|\mathcal{B}]}_{\text{selection}} \\ &= \underbrace{\mathrm{E}[q'^1\left(z'^1 - 1\right)|\mathcal{B}]}_{\text{mutation pressure}} + \underbrace{\mathrm{E}[q'^1 - q^1|\mathcal{B}]}_{\text{selection}}. \end{aligned} \qquad [15]$$

We examine the case of $n = 2$ ancestors. We first calculate $E[q^1|\mathcal{B}]$. Let x_j denote the adopt/reject decision of agent I_j, coded as 0 and 1. Then since $x_1 = 1$, the event $(x_2 = 1, \mathcal{B})$ occurs when (1) $v_1 = V$, or (2) $v_1 = -1$, I_1 is censored, and I_2 is not censored. Also, observe that $q_1 = 0.5(1 + x_2)$. So

$$\begin{aligned} E[q^1|\mathcal{B}] &= 0.5 + 0.5E[x_2|\mathcal{B}] \\ &= 0.5 + 0.5P(x_2 = 1|\mathcal{B}) \\ &= 0.5 + 0.5\frac{P(x_2 = 1, \mathcal{B})}{P(\mathcal{B})} \\ &= 0.5 + 0.5\frac{P(x_2 = 1, \mathcal{B})}{1 - [q(1-p) + (1-q)p]^2(1-\pi)^2} \\ &= 0.5 + 0.5\left\{\frac{qp + (1-q)(1-p) + [q(1-p) + (1-q)p](1-\pi)\pi}{1 - [q(1-p) + (1-q)p]^2(1-\pi)^2}\right\}. \end{aligned} \qquad [16]$$

[†] Higher q, the probability of the good state, increases the absorption buffer (decreases $d^* < 0$), reducing R^L. Higher upside salience β also reduces R^L. Since the probability of a reject is bounded below by zero, a higher buffer weakens the negative marginal effect of upside salience on the probability of ever rejecting. So the cross partial between q and β on R^L is positive.

Table S1. Comparative Statics on Moonshotness

	Levels of moonshotness	
Benchmark vs. Focal project	Initial Level	Increased Level
Benchmark project	$(\overline{\beta}, q, V)$	$(\overline{\beta}, q', V')$
Focal project	(β, q, V)	(β', q', V')

Since rejectors have no causal influence on the descendant population, the descendant population comes entirely from type 1, i.e., $q'^1 = 1$, which is deterministic. So $E[q'^1|\mathcal{B}] = 1$.

We also need to calculate $E[q'^1 z'^1|\mathcal{B}] = E[z'^1|\mathcal{B}]$. The descendant of type 1 has type either 0 or 1, so

$$E[z'^1|\mathcal{B}] = P(z'^1 = 1|\mathcal{B}) \quad = \quad \frac{P(z'^1 = 1, \mathcal{B})}{P(\mathcal{B})}. \qquad [17]$$

The event $z'^1 = 1$ is just the event that I_3 adopts. This occurs iff: (1) both ancestors adopt, and (2) the sole uncensored ancestor has $v = V$, or there are two uncensored ancestors and at least one has $v = V$. Both ancestors adopt iff I_2 adopts, which occurs iff either (a) $v_1 = V$, or (b) I_1 is censored. So the pathways for I_3 to adopt are: (i) $v_1 = V$ (since even if $v_2 = -1$ is not censored, I_3 adopts); (ii) $v_1 = -1$, I_1 censored, $v_2 = -1$, I_2 censored; and (iii) $v_1 = -1$, I_1 censored, $v_2 = V$.

The event $\mathcal{B}$ rules out pathway (ii). So the numerator is

$$\begin{aligned} P(z'^1 = 1, \mathcal{B}) &= qp + (1-q)(1-p) + [q(1-p) + (1-q)p](1-\pi)[qp + (1-q)(1-p)] \\ &= [qp + (1-q)(1-p)]\{1 + [q(1-p) + (1-q)p](1-\pi)\}. \end{aligned}$$

The denominator is

$$P(\mathcal{B}) = 1 - [q(1-p) + (1-q)p]^2(1-\pi)^2.$$

So

$$P(z'^1 = 1|\mathcal{B}) \quad = \quad \frac{[qp + (1-q)(1-p)]\{1 + [q(1-p) + (1-q)p](1-\pi)\}}{1 - [q(1-p) + (1-q)p]^2(1-\pi)^2}. \qquad [18]$$

We have now calculated all the ingredients of the Price equation. Substituting them into (15) gives

$$\begin{aligned} \overline{\widehat{\Delta z}}^{\mathcal{B}} &= \underbrace{\frac{[qp + (1-q)(1-p)]\{1 + [q(1-p) + (1-q)p](1-\pi)\}}{1 - [q(1-p) + (1-q)p]^2(1-\pi)^2} - 1}_{\text{mutation pressure}} \\ &+ \underbrace{\frac{1}{2} - \frac{1}{2}\left\{\frac{qp + (1-q)(1-p) + [q(1-p) + (1-q)p](1-\pi)\pi}{1 - [q(1-p) + (1-q)p]^2(1-\pi)^2}\right\}}_{\text{selection}} \end{aligned} \qquad [19]$$

In this context, in which the ancestral agents may or may not contain a rejector, the censorship rate enters into the decomposition of change from the Price Equation. Since the ratios above are probabilities less than 1, there is again negative mutation pressure and positive selection. The positive selection occurs because only adopters have descendants. The negative mutation pressure occurs because adopters who generate negative payoffs generate descendants who are rejectors.