```
plot_scatterconf (generic function with 1 method)
```

```
Activating project at `~/Documents/proyectos/PessimisticLearning`
```

The co-evolution of pessimism and conformity under environmental uncertainty.

Alejandro Pérez Velilla, Bret Beheim, Paul E. Smaldino.

This is a handout summarizing a forthcoming paper. For internal circulation only.

1. Preliminaries: uncertainty, preferences in risky decision-making and the Kelly model

Risk preferences develop throughout life, a result of personal experience and social learning. In a way, they are the embodiment of the experiences people go through during development. As such, growing up in a safe space ("a bubble") can give people the illusion that the world is less dangerous than it actually is, while growing up in an environment where dangers are evident (or made evident by senior members of the population) can lead to develop behavioral safeguards that operate even when in a safe environment.

In this project we seek to develop a model of how risk preferences are developed, and how environmental uncertainty and social learning impact on such development. In order to do so, we start out with a simple version of Kelly's proportional betting model (Kelly, 1956). In this model, a player starting with wealth w_0 chooses a stake $s \in (0, 1)$, representing the fraction of their wealth they are betting away, and then proceeds to throw a biased coin every time period (for a total of T periods), with every period resulting in a success with probability u (also known as the rate of success) and a failure with probability $1 - u$. If successful, the player increases their wealth proportionally to $1 + s$, and if not successful, their wealth decreases by $1 - s$. At its score, this is a simple model of risky capital investment: the more one places at risk, the more there is to lose and the more there is to win. The less one stakes, the less one can win, but the smaller potential losses will be.

We can write the payoff of a player at time T in the following manner:

$$V_T(s) = w_0 \prod_{t=1}^{T} e^{g_t(s)}$$

with $g_t(s) = \log(1+s)$ if successful, and $g_t(s) = \log(1-s)$ otherwise. It can be shown that in the limit of large T, there is an optimal stake that can be easily calculated and that ensures the best rate of growth for wealth. Known as the Kelly criterion, it can be written as

$$s^* = \begin{cases} 2u-1 & : u > \frac{1}{2} \\ 0 & : u \leq \frac{1}{2} \end{cases}$$

This means that a decision-maker facing a scenario of proportional betting can know the optimal stake if they know the rate of success u. In what remains of this theoretical exercise, we will examine what we think are the main obstacles that learners have to face when trying to be optimal proportional bettors. Throughout life we constantly face decisions that require us to make tradeoffs between risk and reward. The proportional betting model captures this idea with minimal mathematical structure, allowing us to explore how players fare when different forms of uncertainty come into the mix.

We will see that using individual experience and social information can lead to biased information, requiring pessimistic biases (risk aversion) to deal with potentially dangerous optimism that creeps in. Learning from those who came before us can also lead us to choose behaviors that are too risky when we use pure payoff-biased learning and/or when the information we obtain about others exhibits a success bias (successful agents are more visibile, or in the most extreme case, the only visible learning models).

2. Sources of uncertainty in environmental sampling

As story to go along with the argument, we can put ourselves in the shoes of Niko. Niko recently arrived in Liberty City, motivated by the stories of opportunity his cousin Roman wrote about in his letters. Finding himself in a new environment, Niko does not particularly know how to act or what risks to avoid, and he does not want to solely rely on Cousin Roman, who has a weakness for overblown claims. In order for Niko to get a notion of how much it is worth to put his resources at risk, he must have some notion of the city environment's rate of success, u.

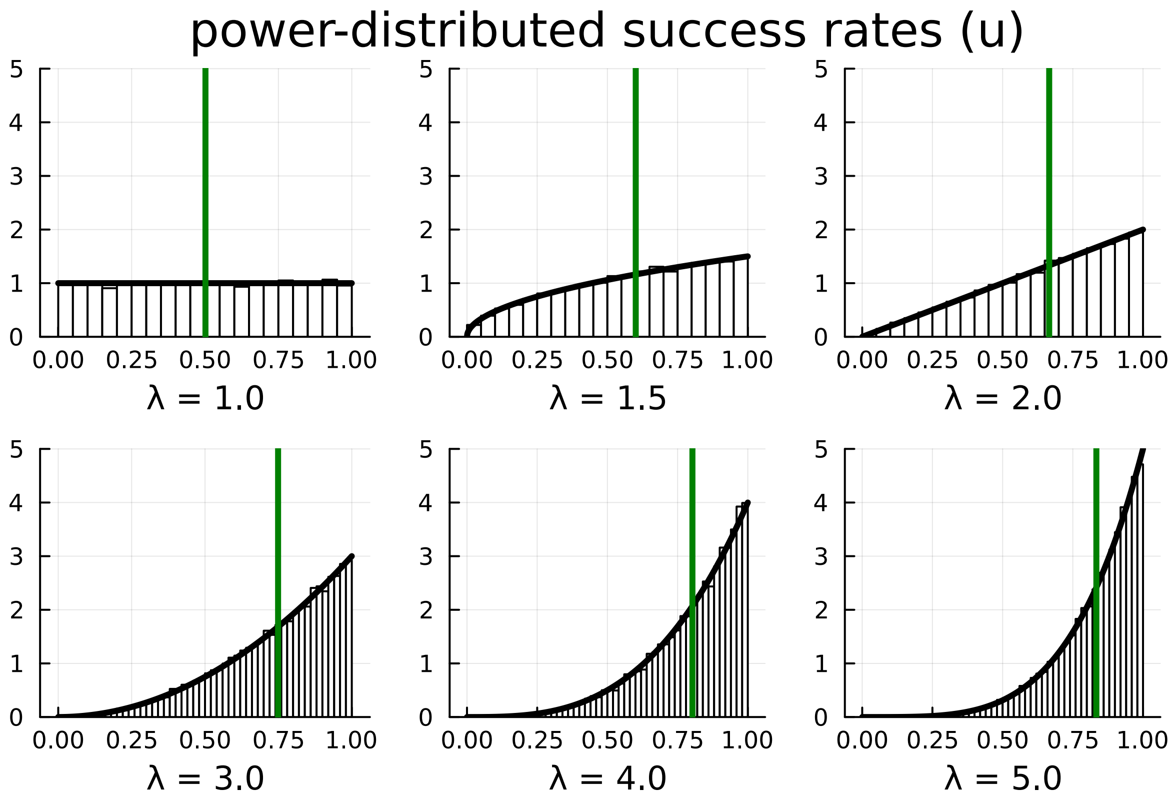
Let us say life in Liberty City is benevolent enough that Niko can sample the environment for a while before settling on a stake. He comes in with enough money to last him a couple weeks, and he does not waste his time. He uses this time to get some sense of how much the city's environment favors (or disfavors) those who take risks, by trying out different activities and noting down his successes and his failures. Since he does not want to rely purely on his individual impressions, he figures that at the end of his testing period he can meet up with other recent arrivals and everyone can share their own impressions of the city. Using this mix of individual sampling and social learning, he should be able to get an estimate of u, which we call $\bar{u}$, which his cognitive machinery transforms into an estimate of the optimal stake one should use ($\bar{s}$).

In doing so, there are several things he must contend with.

Individual learning: imperfect information and random success rates lead to potentially dangerous errors.

Success rates can be random. This means that every time period Niko samples the city, the chance of success he obtains can be different from the last. Thus, every success or failure he gets might have a different success rate generating it, drawn from a common distribution representing the city's stability. Environments where success rates can vary considerably are said to be more *unstable* or *uncertain* than environments where it varies less. If the city is a crime-ridden dystopia governed by a council of mafiosos, Niko's luck might vary significantly on a day to day basis, and he should take care not to put himself in danger. If, on the other hand, the city is a secure place with several safety nets in place, then his chances of success can stay consistently favorable, and he can allow himself to take higher risks.

In general, instead of knowing their chances of success, individuals must sample their environments and try to inferr them. As environments become increasingly unstable, having an accurate estimate requires longer periods of inidividual sampling, which might be infeasible. For this reason, individuals may need to complement their own inferences about the environment with their peers' own estimates, lest they end up staking too little or too much based on luck alone.



We use a Power distribution to model this situation. This means we take the standard Kelly betting model and adds uncertainty in the rate of succes u. Specifically, the *inverse* of the probability of winning a bet follows the Pareto distribution with a minimum value of 1. We will call this quantity V varying between 1 and infinity, and say

$$P(V \leq v) = 1 - (m/v)^{\lambda}$$

where $m = 1$ and $\lambda > 1$.

If V is Pareto-distributed, the random variable $U = 1/V$ follows a Power distribution (which *not* the same thing as a "power law"), and varies between 0 (V is very very large) and m^{-1} (V is close to $m = 1$) (Dallas, 1973):

$$P(U \leq u) = (mu)^{\lambda} = u^{\lambda}$$

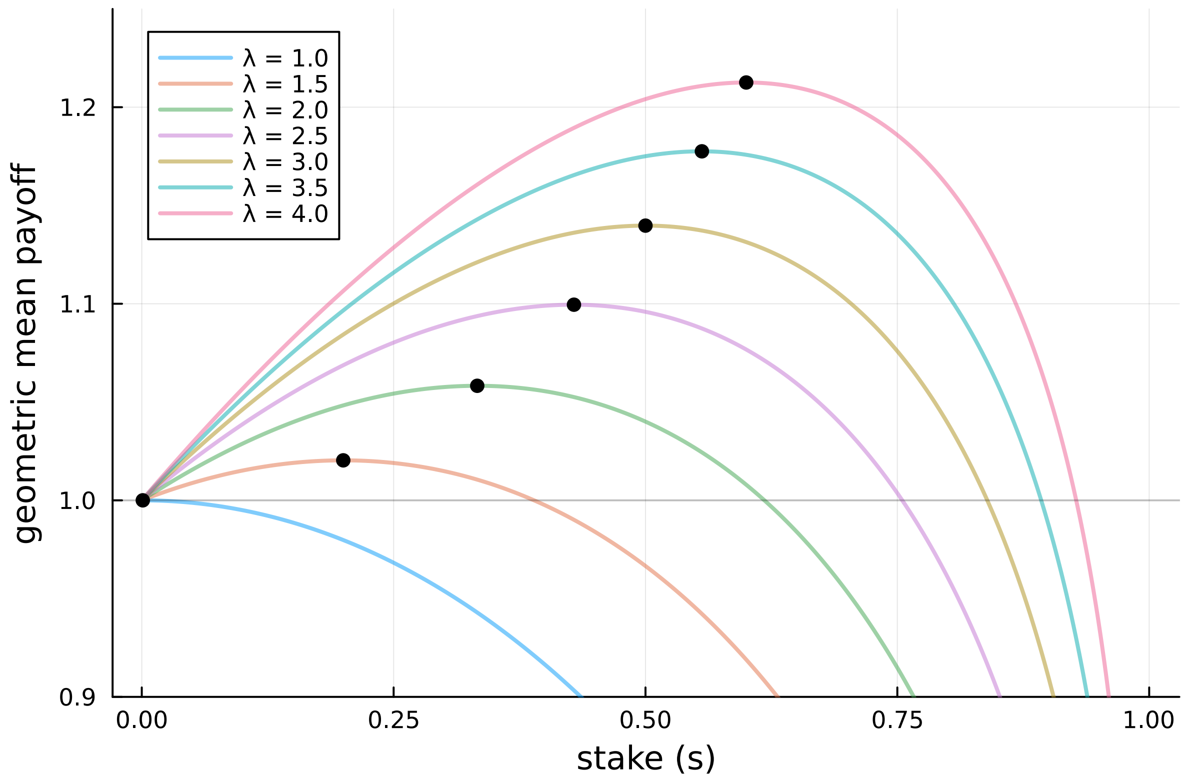
with probability density function

$$f(u) = \lambda u^{\lambda - 1}$$

and an expected value given by

$$E(U) = \frac{\lambda}{\lambda + 1}$$

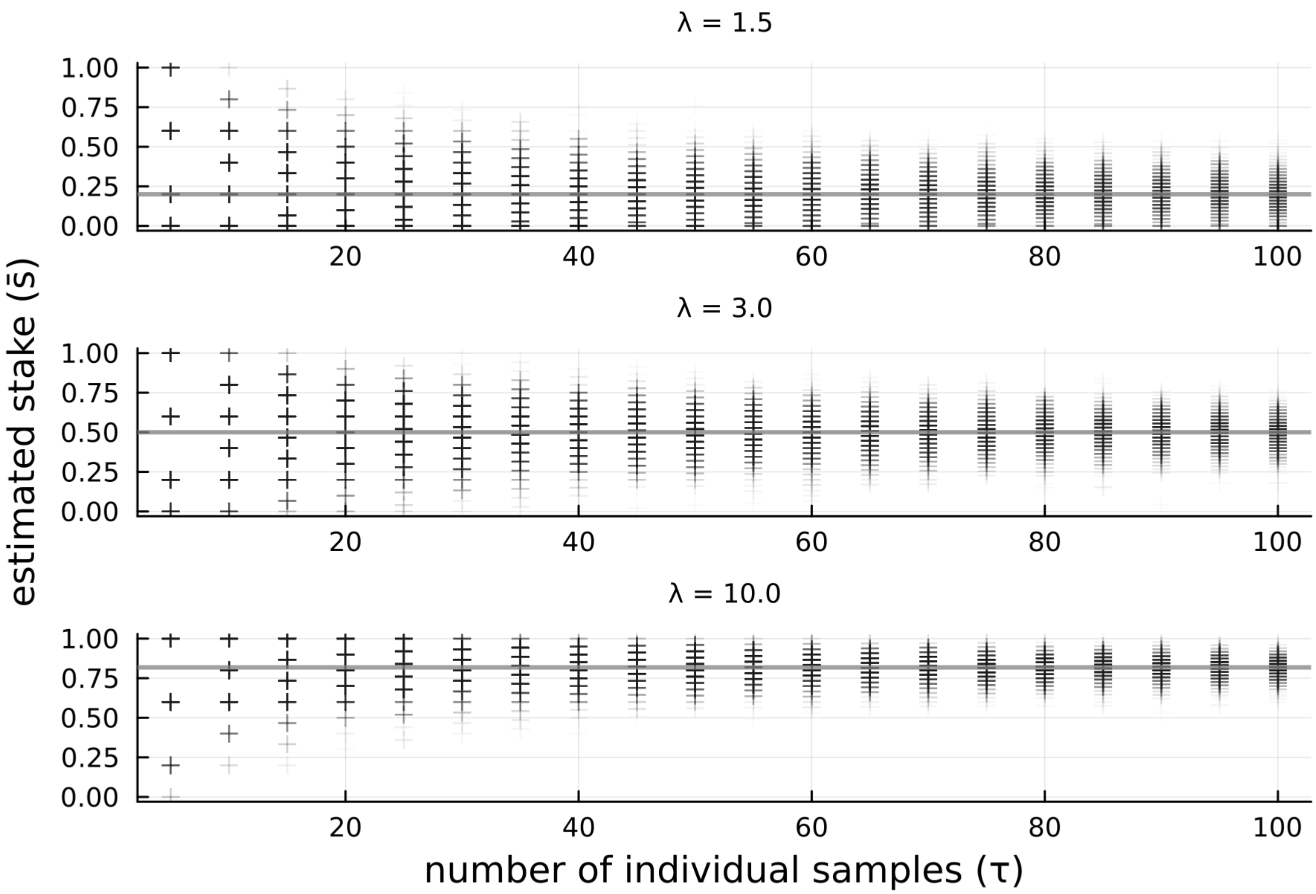
This probability density function is plotted above (smooth line) alongside samples from the distribution (histogram) and the expected value (vertical line) for different values of λ, in order to show how λ works as a parameter controlling the environment's stability/degree of uncertainty. When $\lambda = 1$, the environment is fully uncertain: one can go from having very good fortune to having catatrophically bad luck. As λ grows towards positive infinity, environments grow more consistently favorable, such that the ocassional high failure probability gets more easily outweighted by strings of good fortune.



It can be shown that when success rates are random, then the best value one can use when choosing a stake is the expected success rate, $E(U) = \frac{\lambda}{1+\lambda}$. The figure above shows the mean payoffs for different stakes under different values for λ. The optimal stake for each λ is marked by the black dot on the peak of each curve, and is given by $s^* = 2\frac{\lambda}{1+\lambda} - 1$. The horizontal line marks a growth rate of 1; haivng a payoff below this point means that the individual is incurring in losses rather than gains. It is worth taking not that as λ grows smaller, the region around the optimal stake that supports growth also grows smaller. This means that as environments grow more unstable, there is less of an acceptable margin of error around the stake an individual chooses. **Being overly optimisitc can lead to losses**.

This all means that if Niko knows the city's λ, then he can consistently choose the best stake for his adventures. However, since he does not know it, he must estimate it by sampling the environment before choosing his stake. Over τ time periods, Niko samples the environment by trying out some activity which yields him either a success or a failure, each activity entailing a sample from U. At the end of the τ test periods, he sums up his successes and divides the sum by τ. This yields him an estimate of $E(U)$, which we call $\bar{u}$. Niko then calculates his estimate of the optimal stake, $\bar{s} = 2\bar{u} - 1$.

Below, we simulate many Nikos in order to get an idea of the error rate of such a process, starting at $\tau = 5$ and ending at $\tau = 95$. The grey horizontal line represents the optimal stake given the value of λ. Unsurprisingly, when τ is low the probability that Niko ends up with a very bad estimate is high. The error in estimation decreases as τ grows larger, giving Niko a better chance at getting a good estimate. However, even at large τ, there is still considerable variance around the estimates, which can be dangerous, particularly in more unstable environments where there is less tolerance for bad estimates (see above).

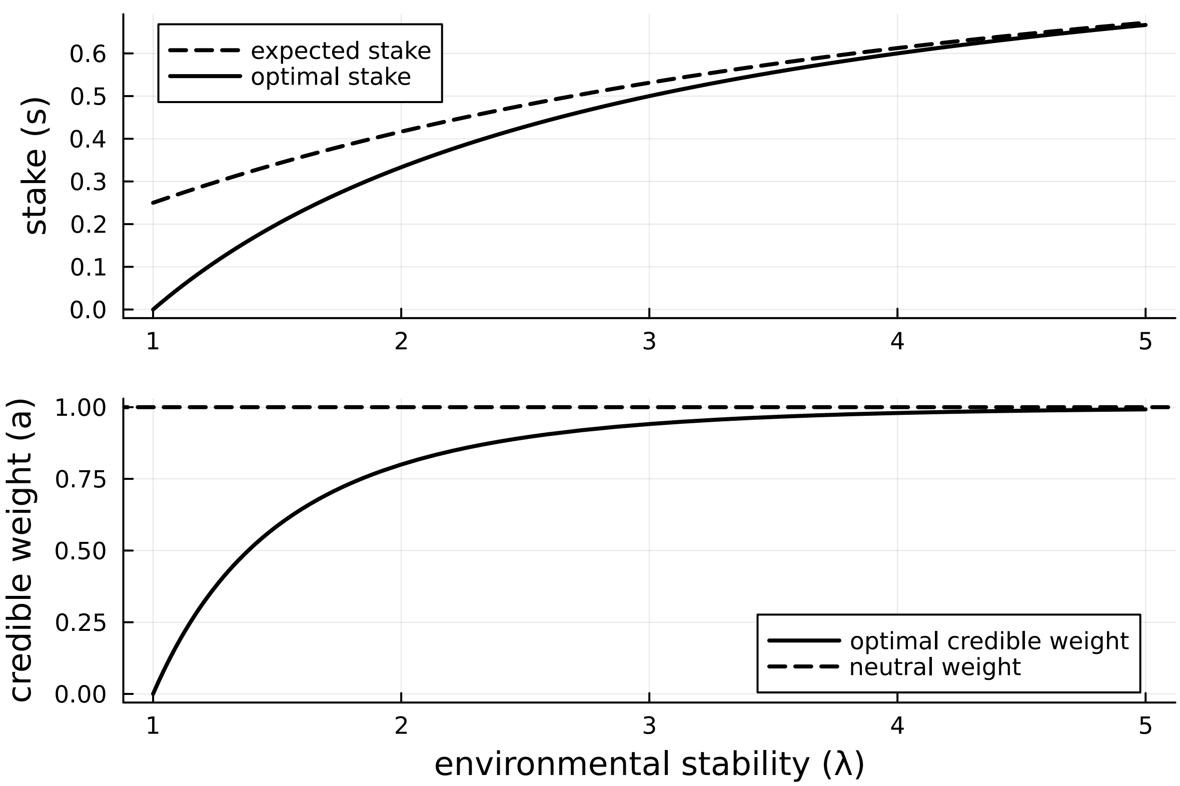


In order to make up for the error rate of individual learning, Niko can turn to social learning in order to get an estimate closer to optimal. We know this from Boyd and Richerson: **when individual learning is prone to error, social learning can save the day**. But there are caveats.

Horizontal transmission: indirect sampling of environmental information leads to unintended optimism.

Socially-sampled information might be biased. When Niko speaks to other immigrants in order to get their impressions, he does not get access to their full experiences, but rather rationalized impressions of them. He now has the daunting task of inferring experiences from behavior.

In general, individuals are not exposed to others' direct experiences. As with Niko, they must use peers' behavior in order to inferr those experiences, and *then* use that information to complement their own sampling. However, behavior is often a non-linear function of experience. Individuals who, based on their individual experiences, judge the environment to be disfavorable for risk-taking will propose not putting anything at stake. On the other hand, individuals who judge the environment as providing favorable chances will propose staking something. In a sufficiently uncertain environment, after individual sampling, there will be a mix of individuals who propose not staking anything and individuals who propose staking something. For a focal individual who wants to use this information to know what to stake, integrating information from these two groups of peers introduces bias in the estimated stakes (see averaging convex functions of random variables and Jensen's inequality).

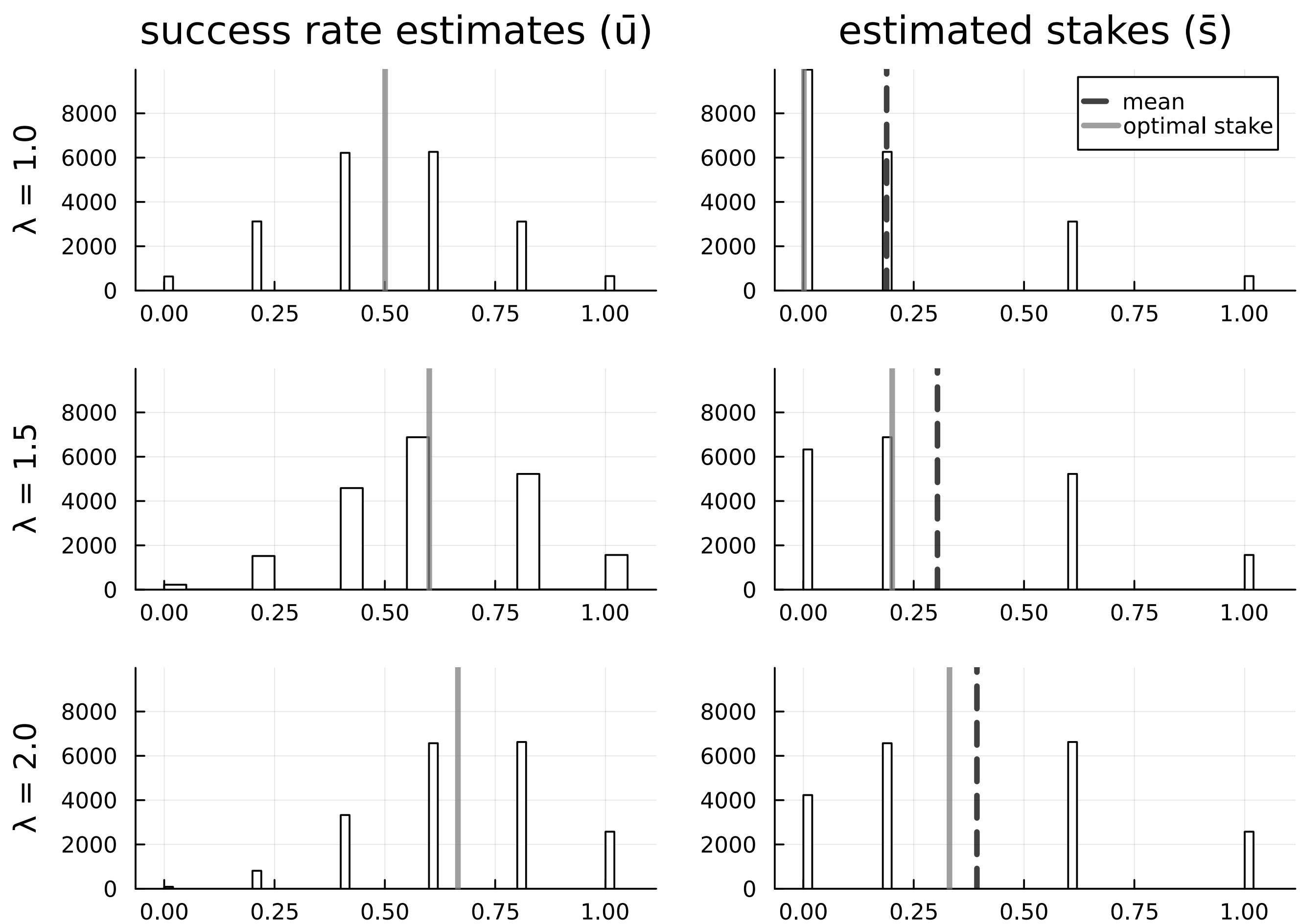


When Niko speaks to another newcomer i, he does not see their $\bar{u}_i$. Rather, he sees their estimate of the optimal stake, $\bar{s}_i$, which is a convex function of $\bar{u}_i$. Jensen's inequality tells us that the distribution of estimated stakes has an expected value that is *higher* than optimal. This means that if Niko tries to hone his estimate by averaging it along with his peers' estimates, he will end up with an overestimate of the optimal stake. In other words, **using social information leads to an optimistic bias**. The above plot (upper pane) shows the difference between $\hat{s} = E(2U - 1)$, the expected stake after averaging social information (dashed line), and $s^* = 2E(U) - 1$, the actual optimal stake. **As environments grow more unstable, this difference becomes more prominent**. This means that social information must be weighted down by a factor $a \in (0, 1)$, which we call a **credible weight**, such that $\alpha\hat{s} = s^*$. This gives us

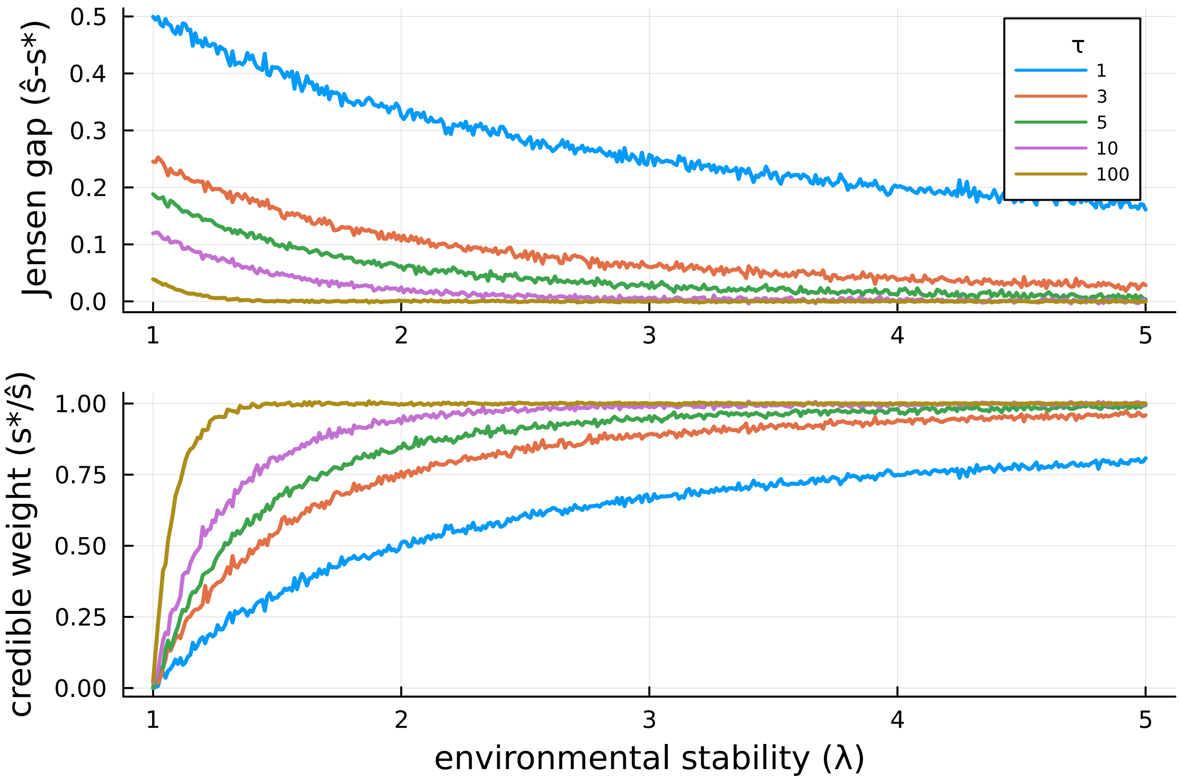
$$a^* = \frac{s^*}{\hat{s}}$$

for the optimal credible weight, which we plot above (lower pane). We can think of α as a measure of an individual's optimal aversion to uncertainty. An a closer to 0 indicates individual should be less confident in obtained estimates, whether they come from individual learning or horizontal transmission, whereas an a of 1 indicates they should exhibit full confidence. It is thus a form of optimal *pessimism*, itself a manifestation of optimal risk aversion.

The above plots assume that each individual draws a value of u from U, which they go on to use in calculating their observable estimate. In practice, however, what each individual draws is an estimate of $E(U)$, with the variance of these draws depending on τ, as we saw above. In order to visualize the dependence on τ, we simulate many immigrants coming into Liberty City, each sampling the city for τ time periods and then sharing their estimated stakes $\bar{s}$ among one another.



The above plots are simulations for $\tau = 5$. The histograms on the left column show the distributions of $\bar{u}$ among immigrants (not observable) while the right column shows the (observable) distribution of $\bar{s}$. The solid line shows the optimal stake for the given value of λ, while the dashed line shows the stake an immigrant like Niko would arrive at if they simply averaged their own estimated stake with the one obtained from their peers. Note how the optimism is induced purely by the averaging process, without a need for individuals to have an optimistic preference themselves.



The above plots show the Jensen gap (the difference $\hat{s} - s^*$ between estimated stakes and optimal stakes) and the optimal credible weight ($a = \frac{s^*}{\hat{s}}$) for different values of τ. It is clear that if test periods are short (τ small) then individuals should apply a more aggresive credible weight, which is to say that they should be more pessimistic. This makes sense, as the quality of social information also depends on the quality of peers' bouts of individual learning. And even when τ is large, more unstable environments require higher degrees of pessimism. **As environments become more unstable, at some point higher degrees of pessimism will be required to mitigate optimisitic biases, regardless of the length of individual learning periods.**

At this point it is clear that being pessimistic can be advantageous through the uncertainty introduced by random success rates and the difference between unobservable experiences and the observable behaviors they lead to. Since staking too much can lead to losses, Niko should be pessimistic in one way or another when dealing with the information he gathers. We now proceed to examine what happens if individuals can lose everything when they hit an absorbing barrier, destroying their ability to bounce back from losses that are too great.

Initial wealth and the looming threat of absorbing barriers: the possibility of ruin intensifies the need for pessimism.

In real life, there are hidden risks that are difficult to account for when choosing a proper course of action. An common manifestation of this is an **absorbing barrier: a "point of no return" for a payoff trajectory**. For example, if payoffs are meant to represent a strategy's reproductive success, then obtaining a payoff of zero in any given generation is an absorbing barrier: having no offspring in one generation means there are no offpsring in any subsequent generations. Similarly, if payoffs represent wealth, then negative wealth growth at any given period can lead to bankruptcy. On the other hand, individuals can also come in with some initial advantage into the betting dynamic, given by an initial level of wealth. Starting off with a higher level of wealth can make it harder to hit that absorbing barrier, by providing more possibilities for bounce-back in case of losses.

Rememeber that at a given time T, the payoff of an individual staking s is

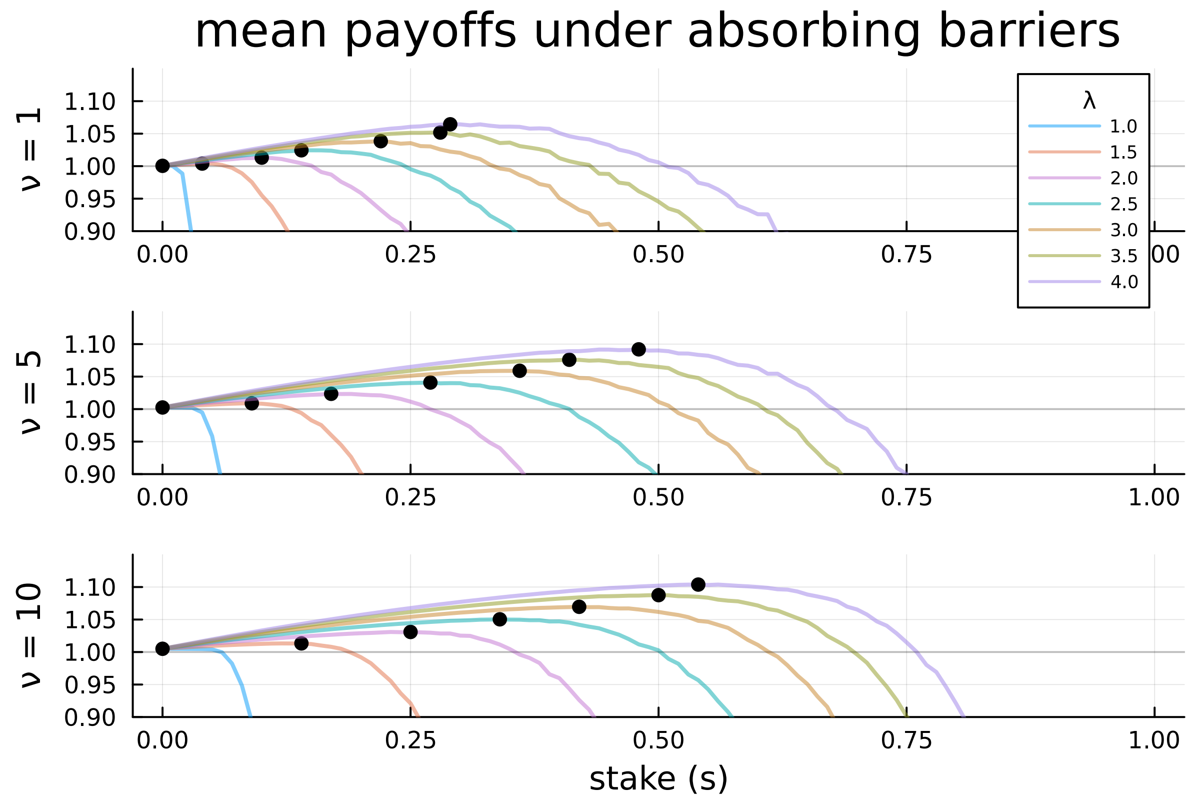
$$V_T(s) = w_0 \prod_{t=1}^{T} e^{g_t(s)}$$

Since we can write w_0 as $w_0 = e^{g_0}$, the above expression is the same as

$$V_T(s) = e^{\sum_{t=0}^{T} g_t(s)} = e^{\nu + \sum_{t=1}^{T} g_t(s)}$$

where $\nu = g_0$ is our measure of initial (log) wealth.

We can then simulate the effects of absorbing barrier and initial wealth by setting an absorbing barrier at $V = 0.1$. This means that if the total payoff at any point of the payoff trajectory becomes lower than 0.1 (which is the same as saying that the agent has lost most of their wealth or soma), then payoffs are set to zero, a state representing total ruin.



The above plot shows what happens to mean payoffs under these new dynamics. The effects of absorbing barriers are dire: optimal stakes are much lower than without barriers, and when individuals start out poor ($\nu = 1$) the optimal stakes are extremely conservative, with very little margin for error. As initial wealth levels increase, so do optimal stakes. However, these are unequivocably lower than the optimal stakes in a dynamic without absorbing barriers, even at very large levels of intial wealth ($\nu = 10$). Since the learning process for estimating success rates is free of risk itself (remember we learn about the environment in relatively risk-free testing periods, like childhood and adolescence), then it cannot represent the effect of absorbing barriers, and thus we can conclude that **absorbing barriers induce a further need for pessimistic weighting, independent of the learning strategy employed to estimate environmental instability.**

Key takeaways #1:

- Individual learning becomes more unreliable as learning periods become shorter and as environments become more unstable. Even long learning periods can lead to significant error when environments are very unstable. Decision-makers thus should be pessimistic about their own estimates after individual learning in order to be on the safe side of the error, as being too optimistic can lead to ruin, while being too pessimistic will only lead to slower payoff growth (or no growth at all in the extreme case of full pessimism, $a = 0$).
- Precautions must be taken when using horizontally-transmitted social information to complement individual learning as environmental stability decreases and environments become more uncertain. An optimistic bias can creep into estimates without any sort of

mechanism explicitly promoting it in individuals' cognitive apparatus. Decision-makers have to use corrective cognitive mechanisms, such as implementing pessimistic credible weights on the information they receive (risk aversion), in order to avoid potentially dangerous unintended optimism.

- The presence of absorbing barriers and the possibility of ruin lead to the need for pessimism regardless of the combination of individual and social learning an individual uses. This makes it even harder to find the appropriate level of pessimism that an individual should use.

Acquiring the right pessimistic cognitive strategy is thus the first-order correction problem of risky decision-making in uncertain environments, applying to individual learning and horizontal transmission. Given that in real life there are always probabilities of ruin for those who take excessive risks, there are clear evolutionary pressures favoring the emergence of a pessimistic bias in decision-making. We argue that this pessimistic bias is more likely to arise from cultural evolutionary processes rather than through genetic evolution, as human environments can change from relatively stable to relative unstable in the span of mere generations, prompting decision-makers to use learning as an adaptive tool. Once test periods are finished, learning is not necessarily finished: as we observe the successes, losses and ruin of others around us, more information is revealed about the environment than that which we can get during sheltered testing periods. For example, a sensibility to observed ruin events can use information about visible catastrophic failures without having to be a victim of one. By using this information to adjust pessimistic credible weights as a function of how frequent ruin is in the environment, decision-makers could arrive at the optimal degree of pessimism, although some of them will inevitably fail.

The kind of learning described until now is good for when environmental conditions change from one generation to another. But when environments maintain the same level of initial wealth and environmental stability for many generations, learners need not expose themselves to the possible errors of individual learning and horizontal transmission. Instead, vertical and oblique transmission can cue learners into what has been working for the past generations, who already performed in the environment. Thus, learners can rely on culture as an adaptive tool. This leads to the second order problem of risky decision-making in uncertain environments, which we examine below.

2. Sources of uncertainty in social learning (vertical and oblique transmission)

Niko knows there are many unknowns about the information he has managed to get about life in Liberty City, and that he should be careful. But being too careful might lead him to passing on actual good opportunities. Luckily for him, his is not the first wave of immigrants to arrive at the metropolis. He sets out to find those who came before him and hear their stories of success and their advice on how pessimistic to be.

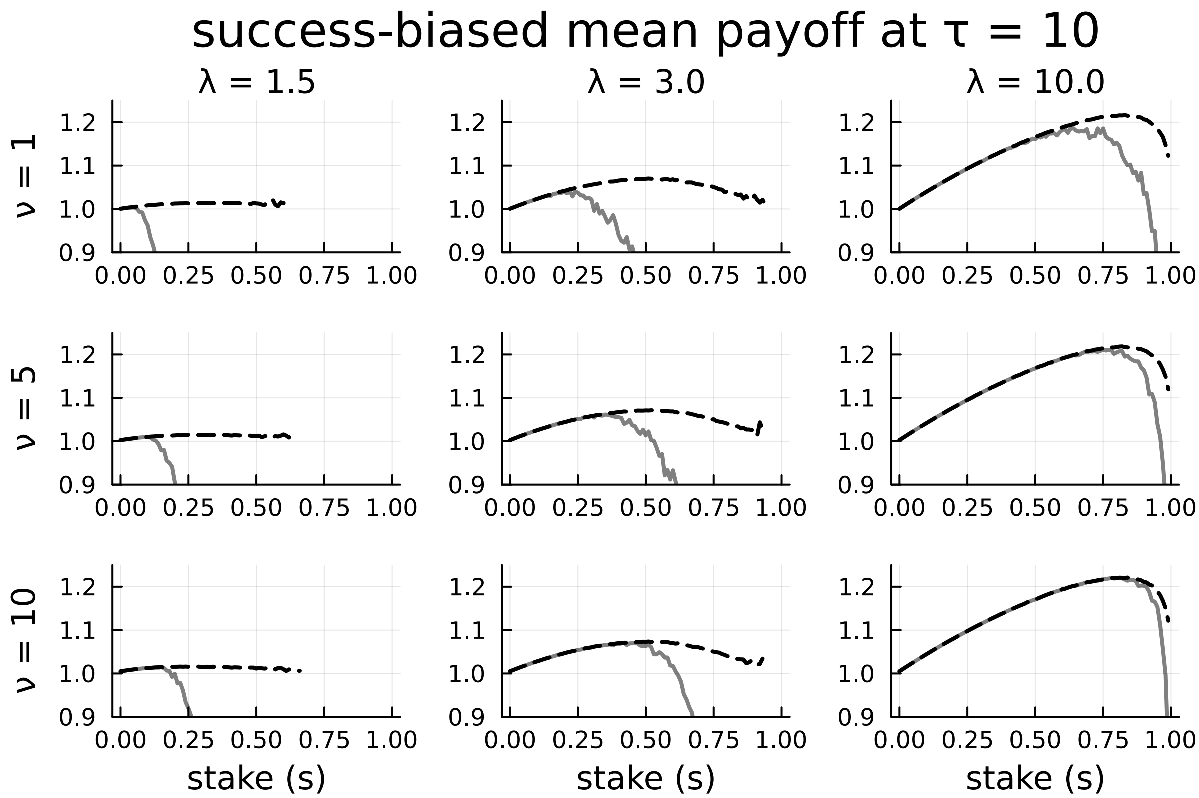
Niko speaks to Uncle Yuri, the owner of a successful deli who tells him how he put all of his savings behind his then-nascent business. On TV, he watches broadcasts of immigrant success stories like that of millionaire Edna, who says she found a welcoming city full of opportunity. This fills Niko with a sense of confidence; surely it is not necessary to be that pessimistic! In paying attention to these success stories, Niko inadvertedly exposes himself to danger yet again, as he fails to consider information about those who failed and experienced ruin. While Uncle Yuri's and Edna's successes are commendable, they are also more visible and salient than the failures of those who might have been using equally optimistic credible weights.

Success-biased social sampling hides failures, leading to dangerous optimism.

Success bias is not the same as payoff bias. While payoff bias looks at a distribution of strategies alongside their payoffs and preferentially chooses among the strategies with the highest associated payoffs, **success bias is about visibility: agents that did well enough to be thought of a successful are those who will be considered as possible targets for learning**. Thus, success bias works at the level of the learning pool, and only thereafter can another learning strategy such as unbiased learning or payoff bias act upon the pool of possible targets.

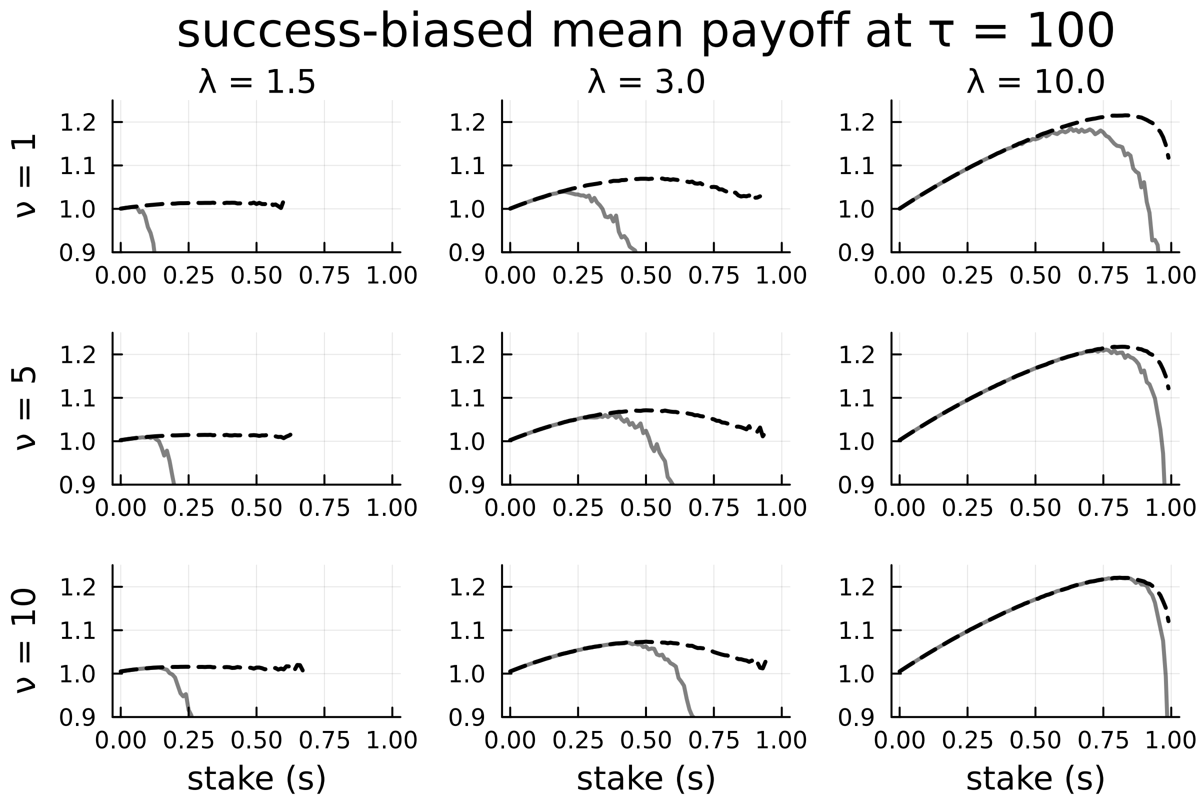
When successes are disproportionally visible (or, in the most extreme case, the only visible behavior), then the optimal degree of pessimism can actually appear suboptimal, while dangerously-optimistic weighting can seem like the highest-paying solution. This happens because the higher variability of optimistic weighting leads to long tails in payoffs, which under success-biased sampling can end up generating skewed estimates of the average payoff of a weighting strategy. Success bias is likely more common today than in days past, due to mass media that overblows the visibility of highly successful individuals, to the point of having ad campaigns that have recognized individuals vouch for behaviors that are completely alien to the source of their success. This is an example of how, just like the estimation error in the previous section, an optimistic bias can arise without there being mechanisms generating it at the level of an individual's cognition. Instead, success bias can arise purely from external processes, such as market competition between mass media outlets.

When looking at previous generations in order to choose a the right degree of pessimism, success bias can paint a distorted picture of what works best. This is due to the fact that the least pessimistic strategies also have the highest payoff variance, so that the few non-pessimistic individuals who are lucky enough to succeed can also happen to do so with very high payoffs. Thus, the average surviving payoff of a given level of pessimism can look very different than the actual average payoff, which includes the (zero) payoffs of ruined individuals.



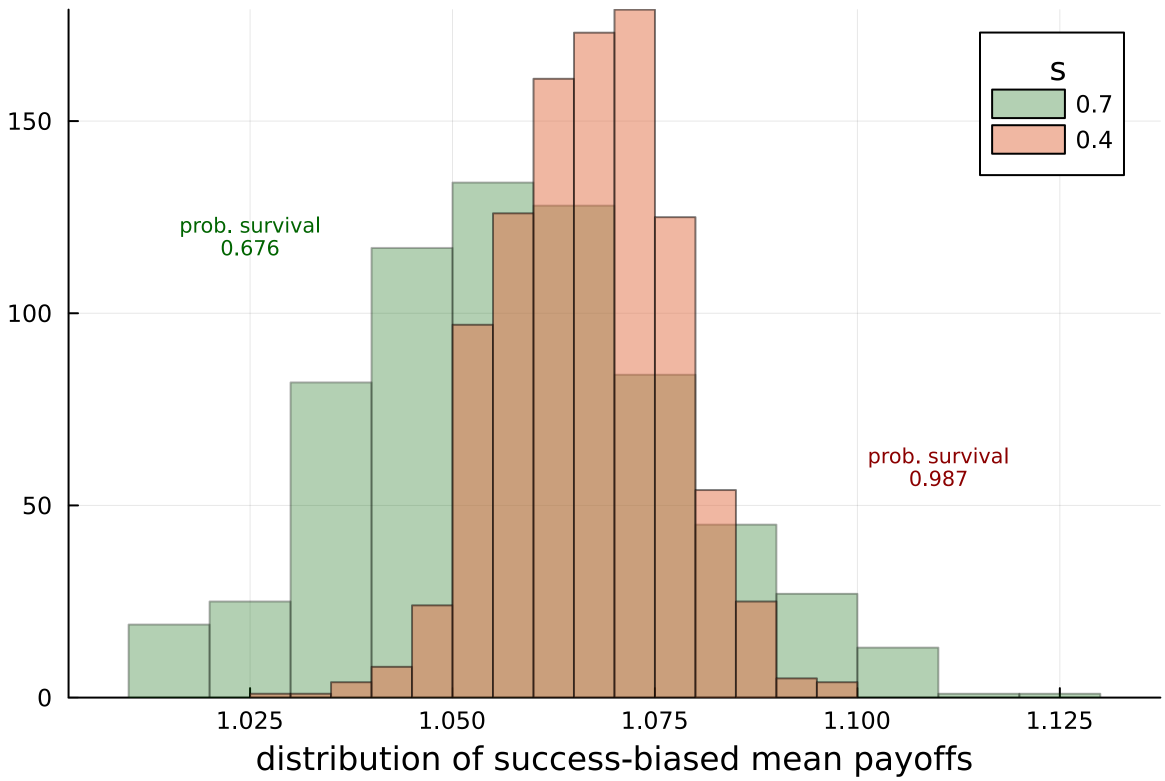
In the above plot, bets are simulated for values of s going from 0 to 1, in steps of 0.01. The dashed line shows the success-biased mean payoff at a given s (in which only payoffs greater than 0 are considered), while the solid grey line shows the full mean payoff. It is evident that in most cases these two payoffs will diverge at some point as s goes from 0 to 1, which is meant to show how much success bias can distort the payoff landscape for learners.

The grey solid line is what decision-makers want to optimize for, so its considerable difference from what is actually visible (the dashed line) poses a significant challenge, except perhaps when learner populations start very rich and environments are very stable. Very little changes by making individual learning periods longer, as can be seen in the plot below. How, then, can the optimal strategy be recovered from information that has been sifted through a success-biased filter?



Payoff-biased social learning also leads to optimism due to the higher variance of less pessimistic strategies.

Even in the absence of success bias, payoff bias on its own can still pose problems. What appears as a handy learning strategy where we focus on the highest observed payoffs and adopt the strategies associated to them can lead us to make fatal mistakes when the strategies considered have to do with risk management. However, it is likely that success bias and payoff bias often occur together.



The above figure is an illustration of how payoff bias can make agents fall for one of the classic blunders. Plotted for $\lambda = 3$, $\nu = 5$ and $\tau = 100$, the figure shows the payoffs of 500 agents who go through the proportional betting game for 2000 time periods. This procedure is done for $s \in [0.4, 0.7]$. It is evident from looking at the histograms that 0.4 is the better-performing stake, both in terms of modal payoff and probability of survival (which is just the complement of the probability of ruin). However, the higher variance of the less pessimistic stake leads to a few payoffs that are considerably higher than the more pessimistic strategy. This right tail of high payoffs will be especially salient to payoff-biased learners, nudging them towards suboptimal (even dangerous) levels of pessimism.

Paying special attention to the highest payoffs of the previous generation leads decision-makers to honing in on the tails of the observed payoff distribution. This can lead them to preferentially choose overly-optimisitic strategies that worked well for a few individuals, but that will lead to bad payoffs or even ruin for most individuals who use them. Both with success-biased social sampling and without, only conformity/frequency-biased social learning, in which information about frequency is used to penalize rare successes before searching for the highest payoffs, will steer decision-makers away from disaster, while still managing to use payoff information strategically.

Frequency biases can recover optimal pessimism from success-biased model pools, while reaping the rewards from payoff biased strategies.

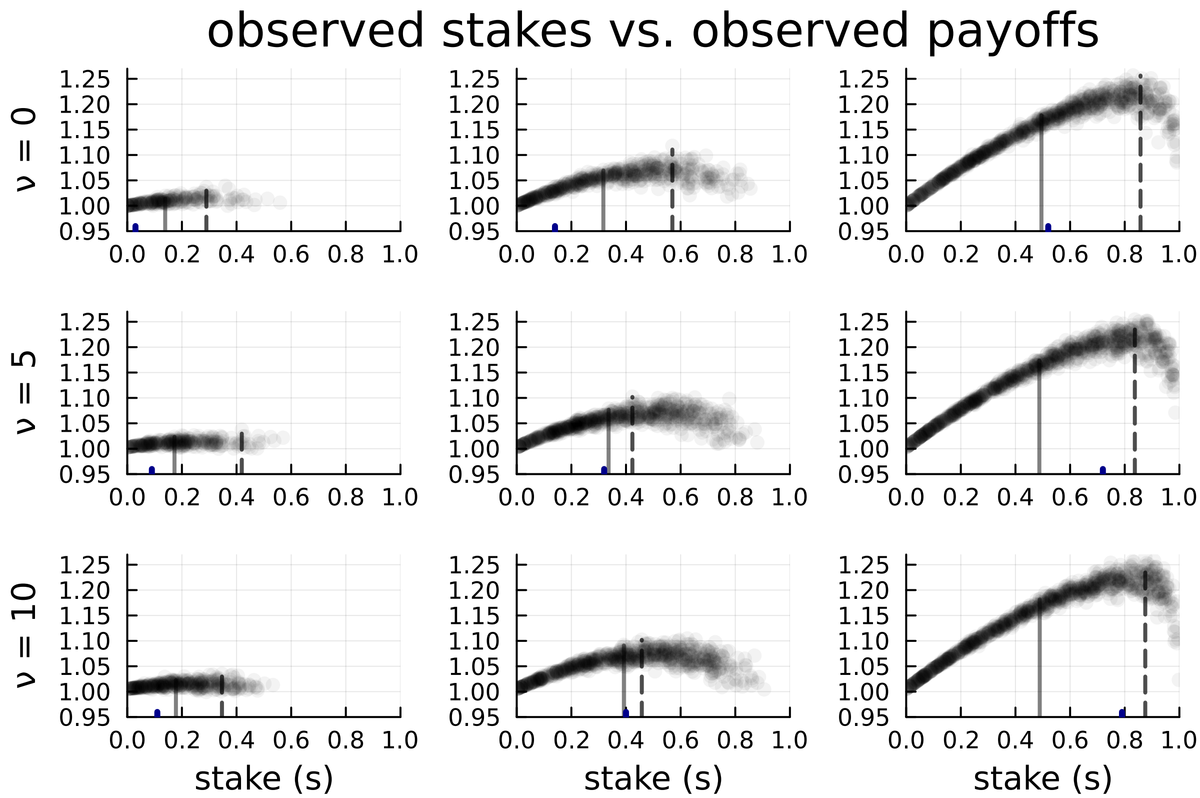
When looking at the previous generation and observing their strategies and payoffs, one can be conformist by looking at the median strategy present in the population. If a decision-maker is fully conformist, then they would immediately choose the median strategy. In a fortunate turn of events, this works well in the presence of success bias: since the riskier strategies are less represented in the learning model pool, being conformist will avoid these. But perhaps there are circumstances in which incorporating payoff information might still be useful. Must we choose between one strategy or the other? Not necessarily. For a set of observed payoffs $v \in V$, each associated to an observed strategy $s \in S$, we can write down a hybrid frequency-payoff biased learning rule like so:

$$w(s, v) = K_S(s) \cdot v$$

where $w \in W$ is then a frequency-weighted payoff, and

$$K_S(s) = \exp\left(\frac{(s - M[S])^2}{\zeta}\right)$$

is a weight determining the effect of frequency bias on payoffs, such that payoffs associated with deviations from the median observed strategy $M[S]$ are penalized. The degree to which they are penalized depends on ζ, which is the parameter controlling frequency bias. When ζ approaches 0 from the right, learners approach full conformity, choosing purely based on how close a strategy is to the median observed strategy. On the other hand, as ζ grows large, all weights approach 1 regardless of distance from the median. This means that if one chooses the highest payoff from the set of weighted payoffs W, one can get pure conformity, pure payoff bias, or anything in between by varying the ζ parameter.



The above figure shows scatterplots representing what a learner sees when they see the previous generation of decision-makers. For each plot, 1000 decision-makers are simulated using strategies drawn from a Beta(1, 1) distribution. The focal agent then sees both the weight they used as well as the associated payoff, for all agents who managed to survive. The dashed line is the stake a learner using pure payoff bias would choose: the one associated to the highest observable payoff. The solid line represents the median stake among survivors: the s that a purely frequency-biased learner would end up choosing. Between these two, ζ can be tweaked to obtain a form of hybrid frequency-payoff bias. The blue notch on the x-axis of the plots is the optimal stake, given λ and ν. Thus, we arrive at the following conclusions:

- In very unstable environments (here given by $\lambda = 1.5$), one should be fully conformist, regardless of initial wealth. Even by being conformist, a degree of optimism cannot be fully avoided, but conformity does help mitigate the dangers of overconfidence. In environments like these, one should be both conformist and extra pessimistic regardless of initial wealth level, although higher initial wealth lessens the need for extra pessimism.
- As environments become more stable, but still fluctuate considerably (here given by $\lambda = 3$), full conformity gets closer to optimality, especially as initial wealth increases. Everyone is best by being conformist here, but the wealthy certainly do best, while the downtrodden should still be extra pessimistic.
- As environments stabilize further (here given by $\lambda = 10$), the optimal learning strategy will be an implementation of hybrid frequency-payoff bias. The degree of frequency bias depends on decision-makers' intial wealth. If the learner population starts out with scarce resources, then

full conformity performs close to optimality. At the other end of the spectrum, if the learner population starts out wealthy, then a lower degree of conformity should be kept, as the optimal stake approaches the one that pure payoff bias would choose.

Key takeaways #2: When using social learning to find the right degree of risk aversion by looking at the performance of previous generations, individuals are prone to a second-order correction problem, analogous to the first-order problem described above. Using payoff information can lead to a biased estimate of what pessimistic strategy works best, as the attractiveness of a strategy will be a non-linear (convex) function of its payoffs. Pure payoff bias can lead individuals to choose strategies that worked very well for the few, but spelled disaster for the many. This problem grows more dire in the presence of low initial wealth and as environments grow more unstable and uncertain. Incorporating frequency information through a cognitive strategy of conformist bias (weighting the payoffs of strategies by their distance to the median surviving strategy) is a possible solution to this problem.

These results imply that natural selection will favor the cultural evolution of frequency-biased learning, especially in environments that have a continued generational history of instability. Only in the combination of highly initially wealth populations in particularly stable environments does conformity start losing its edge, with pure payoff bias leading to the best performance. Risk aversion should always be present in one way or another, although it should decrease as environments become more stable. Thus, pessimism/risk aversion should co-occur with conformity, and this association should be stronger among decision-makers of lower initial wealth. If decision-makers come from wealthier backgrounds, then this association should weaken in stable environments, but remain strong in unstable environments.