Embarrassment Aversion
(Draft Version)

Rick Harbaugh*

October 6, 2023

Abstract

The social psychology literatures on self-esteem, achievement motivation, and self-handicapping have long shown that a fear of embarrassment affects behavior involving risk. We reformulate these early insights as a Bayesian model of embarrassment aversion – or risk aversion with respect to estimated skill. Embarrassment aversion can rationalize the assumptions of classic models, and in particular implies choices consistent with the prospect theory anomalies of loss aversion, probability weighting, and framing. Loss aversion arises because losing any gamble, even a friendly bet with little or no money at stake, reflects poorly on the decision maker's skill. Probability weighting emerges because winning a gamble with a low probability of success is a strong signal of skill, while losing a gamble with a high probability of success is a strong signal of incompetence. Framing matters by altering equilibrium beliefs, such as daring someone to take a risk rather than admit to a lack of skill. Since the predictions of embarrassment aversion depend on the social and information context, the theories make diverging predictions in specific situations. D81; D82; C92; G11

*Kelley School of Business, Indiana University, Bloomington, IN (email: riharbau@indiana.edu). Early versions of this paper have circulated under different titles. I thank David Bell, Roland Benabou, Tom Borcherding, Vince Crawford, Nandini Gupta, Bill Harbaugh, Wayne Harbaugh, Ron Harstad, Tatiana Kornienko, Ricky Lam, Marilyn Pease, Lan Zhang, and seminar participants at the Claremont Colleges, DePaul University, Emory University, IUPUI, Princeton University, UNC/Duke, the Behavioral Research Council, the Econometric Society Winter Meetings, the Stony Brook Game Theory Conference, the Midwest Theory Meetings, and the Berlin Status and Social Image Workshop.

Most risky decisions involve both skill and chance. Success brings both material gain and an enhanced reputation for skill, while failure is doubly unfortunate. The manager of a successful project earns monetary rewards and future opportunities, while the manager of a failed project loses both financially and reputationally. An investor who picks a successful stock enjoys the esteem of friends and family, while an investor who chooses poorly looks like a foolish loser.

Classic social psychology theories including self-esteem and impression management (James, 1890; Goffman, 1967), achievement motivation (Atkinson, 1957), and self-handicapping (Jones and Berglas, 1978) have long recognized that people choose among risky alternatives in part to avoid looking unskilled. While these early theories use reduced form approaches that do not model information flows, the career concerns literature formally analyzes how observer estimates of a manager's skill are updated in risky environments (Holmstrom, 1982/1999, 2016). The literature considers both performance skill that helps a manager succeed at a given project, and evaluation skill that helps a manager choose a project with a higher chance of success.[1]

Based on the career concerns approach and the general social signaling literature starting with Spence (1973), we revisit the early psychology literature on the fear of looking unskilled. We assume that people are "embarrassment averse" in the same pattern typically assumed for risk aversion regarding wealth – they dislike uncertainty over perceived skill and they are downside risk averse in that they particularly dislike looking inept. Such risk aversion regarding perceived skill could reflect damage to future earnings, a particular concern for job security (Chevalier and Ellison, 1999), fear of lost status, or just a personal preference. We then adapt the notion of a risk premium to derive embarrassment premia for gambles based on equilibrium beliefs and posterior skill distributions.

We show that insights and experimental results from the early social psychology literature are generated by this combination of Bayesian updating and embarrassment aversion. Consistent with the self-esteem and impression-management literatures, losing a gamble implies there is a good chance that the decision maker underperformed (performance skill) or misevaluated the odds (evaluation skill). Either case reflects poorly on the decision maker's skill, so decision makers may choose whether to gamble based on how the outcome and the choice itself affects impressions of their skill. Consistent with the insight from the achievement motivation literature that there is "little embarrassment in failing" at difficult tasks and a great "sense of humiliation" in failing at easy tasks (Atkinson, 1957), decision makers may choose long-shots over sure-things to limit embarrassment. Indeed, as in the self-handicapping literature, deliberately undermining the odds of success can mitigate the embarrassment, even when the choice to self-handicap is itself a bad signal.

We then ask what behavioral anomalies will appear to result if people are concerned with avoiding embarrassment, but are modeled as caring only about immediate monetary outcomes. Adapting the notion of a risk premium to estimated skill, we derive embarrassment premia for gambles based on equilibrium beliefs and posterior skill distributions. Using this approach we find that behavior will appear to deviate from expected utility theory in accordance with the canonical behavioral anomalies of loss aversion, probability weighting, and framing as formalized in prospect theory (Kahneman and

[1] Performance skill offers insight into phenomena from rat race career incentives (Holmstrom, 1982/1999) to corporate conformism (Zwiebel, 1995), and evaluation skill into phenomena from distorted investment decisions (Holmstrom, 1982/1999) to political correctness (Morris, 2001).

Tversky, 1979; Tversky and Kahneman, 1981, 1982; Kahneman, 2002). Hence applying the formal approach of the career concerns literature to early approaches from social psychology can provide a rationale for prospect theory based on social concerns rather than perceptual biases.

Regarding loss aversion, since losing a gamble reflects poorly on the decision maker's skill, embarrassment aversion makes her more averse to gambling than pure risk aversion regarding monetary payoffs would predict.[2] Moreover, since losing even a friendly bet with little or no money at stake is embarrassing, this effect becomes relatively more important as the stakes of the gamble become smaller.[3] Hence the utility function when measured based on wealth alone will not be locally linear (Pratt, 1964), but will appear to have a kink at the status quo (indeed, a discontinuity if measured sufficiently finely) as in the standard loss aversion model (Kahneman and Tversky, 1979).

Regarding probability weighting, failure at a long-shot is common but has little effect on perceived skill since both skilled and unskilled decision makers usually fail, while failure at a sure-thing is rare but far more embarrassing since a person who fails is probably unskilled. We find that likely but less embarrassing losses offer higher expected utility than unlikely but humiliating losses. Hence people will appear to overweight low probabilities of success by being less afraid when success is unlikely, and to underweight high probabilities of success by being more afraid when success is likely (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992).

For performance skill, such probability weighting is strengthened if the decision maker has private information about her own skill, in which case failure to take a gamble can be seen as revealing a lack of confidence. Long-shots offer little embarrassment from losing, so it is often worthwhile to take the gamble rather than admit low skill by not trying. For evaluation skill, probability weighting is strengthened if the outcome of a rejected gamble is still observable. A good outcome then indicates that the decision maker failed to recognize that the gamble had better than expected odds. For long-shots this potential embarrassment from not taking a gamble that wins is worse than the potential embarrassment from taking a gamble that loses.

Regarding framing, multiple equilibria often coexist depending on whether the observer expects the gamble to be taken, so a decision maker might be "dared" into taking a chance rather than admit a lack of skill by not trying. Depending on whether losing or winning is portrayed as the reference point, the decision maker may expect that refusing to take a chance will be viewed neutrally or negatively.[4] Consistent with framing effects from prospect theory, such beliefs predict risk aversion when the outcomes are presented as a gain relative to the reference point, and risk lovingness when the same outcomes are presented as a loss relative to the reference point (Tversky and Kahneman, 1981).[5]

[2] Not all situations fit embarrassment aversion. An employee given an output quota (Weitzman, 1980) or a contestant facing an opponent (Charness, Rustichini, and van de Ven, 2018) may benefit from being underestimated. And a manager hoping for promotion may gain from a more variable skill estimate (Holmstrom and Costa, 1986).

[3] This is consistent with Robert Schlaifer's (1969, p.161) suggestion that in some cases "nonmonetary consequences" of losing may explain high risk premia for small gambles.

[4] In the so-called "Asian flu" example with the lives of 600 people at stake, in one formulation not taking a chance is described as "200 people will be saved" and in the other as "400 people will die" (Tversky and Kahneman, 1981).

[5] The existence of multiple equilibria implies a role for cultural factors. For instance, men might take more risks in particular situations (Brad Barber and Terrance Odean, 2001; Catherine Eckel and Phillip Grossman, forthcoming; Rachel Croson and Uri Gneezy, 2009) because if men are expected to gamble then the negative inference from not taking a gamble is larger for men, making the beliefs self-fulfilling. Consistent with a multiple-equilibrium perspective, decision maker

Since the predictions of embarrassment aversion are sensitive to the social and information context, they depend on this context and can therefore differ from prospect theory. First, embarrassment aversion predicts that prospect theory behaviors should be stronger the larger is the skill component.[6] Second, the behaviors should be stronger the greater is the potential observability of the outcome. Third, loss aversion should not disappear in the limit as monetary stakes go to zero as the standard kinked utility function of prospect theory implies. Fourth, if the observer is uncertain whether a given gamble is a long-shot or sure-thing the observer's skill estimate given winning or losing is less affected by the type of gamble. For sufficient uncertainty, the embarrassment premium is then higher for a long-shot so the preference for long-shots over sure-things is reversed.

1 Embarrassment Aversion

We start by analyzing performance skill for the simplest case where the decision maker does not have any private information about their ability to succeed. Based on results from this model, we then consider asymmetric information where the decision maker has a noisy signal of their likely success, so the choice to gamble or not can itself be revealing as part of a signaling game. For performance skill, the choice to not gamble is indicative of a lack of decision maker confidence in their ability. For evaluation skill, the choice reflects their evaluation of the likely success of the gamble, so the outcome of the gamble is a noisy indication of the decision maker's evaluation skill. We then extend this case to the outcome of a gamble being observable even if it is not chosen, implying that not only is taking a gamble that fails embarrassing, but so is failing to take a gamble that succeeds.

1.1 Performance Skill: No Private Information

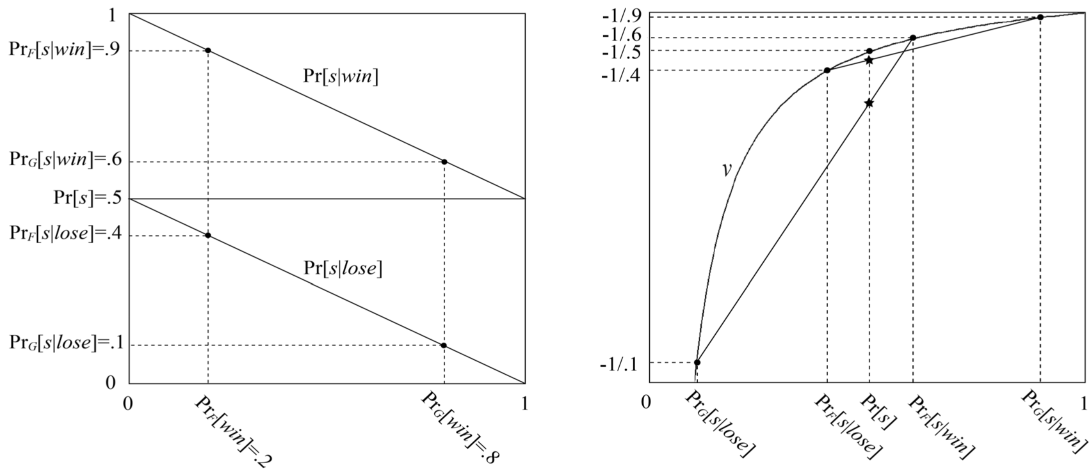
A decision maker faces a gamble with two monetary outcomes $x \in \{win, lose\}$ at opportunity cost or price z where $win > lose$. The decision maker is either skilled "s" or unskilled "u" where a skilled type is more likely to win, $\Pr[win|s] > \Pr[win|u]$. The decision maker does not know their ability $a \in \{s, u\}$. Define the gamble by the distribution $F(x, a)$ which has full support.

Decision maker utility U is determined by both wealth y and by estimated skill μ to an observer who knows the gamble's odds and sees its outcome. This concern for looking skilled could reflect instrumental factors as in the career concerns literature or even an internalized preference. Assuming $U(y, \mu)$ is separable in y and μ and increasing and continuous in y, it is without loss of generality for our results based on the skill estimate component to assume quasilinearity, $U(y, \mu) = y + v(\mu)$.[7]

behavior often varies depending on what aspects of identity are made salient (Akerlof and Kranton, 2010).

[6] Classic prospect theory experiments involved hypothetical games where it is unclear whether subjects should imagine a real world environment with a skill component or not. It is difficult to fully exclude skill from experiments since subjects vary in their ability to understand instructions and even in their ability to choose experiments with better potential payoffs. Hence a subject who does well may justifiably feel that their choices reflect well on them, and a subject who does poorly may understandably feel embarrassed.

[7] If the concern for μ is purely instrumental for its effect on future wealth, separability is most appropriate for small financial stakes where utility in wealth is locally linear in the monetary outcome.



a) Posterior Skill Estimates for Long-Shot F and Sure-Thing G

b) Skill Estimate Component of Utility $v = -1/\Pr[s|x]$

Figure 1: Performance Skill: Posterior Skill Estimates and Expected Utility

Conditional on winning or losing, the estimated skills μ are

$$\begin{aligned}
\Pr[s|win] &= \frac{\Pr[win,s]}{\Pr[win]} = \Pr[s] + \frac{\Pr[win,s] - \Pr[s]\Pr[win]}{\Pr[win]} \\
&= \Pr[s] + \frac{\Pr[win|s]\Pr[s] - \Pr[s]\left(\Pr[win|s]\Pr[s] + \Pr[win|u]\Pr[u]\right)}{\Pr[win]} \\
&= \Pr[s] + \frac{\Pr[win|s]\Pr[s] - (1-\Pr[u])\left(\Pr[win|s]\Pr[s] + \Pr[win|u]\Pr[u]\right)}{\Pr[win]} \\
&= \Pr[s] + \frac{\Pr[win|s]\Pr[s]\Pr[u] - (1-\Pr[u])\left(\Pr[win|u]\Pr[u]\right)}{\Pr[win]} \\
&= \Pr[s] + \frac{\Pr[win|s] - \Pr[win|u]}{\Pr[win]}\Pr[s]\Pr[u]
\end{aligned} \tag{1}$$

and

$$\Pr[s|lose] = \Pr[s] - \frac{\Pr[lose|u] - \Pr[lose|s]}{\Pr[lose]}\Pr[s]\Pr[u]. \tag{2}$$

Notice the skill estimate is updated more strongly the more unlikely is the outcome x, the larger is the "skill gap" $\Delta \equiv \Pr[win|s] - \Pr[win|u] = \Pr[lose|u] - \Pr[lose|s] > 0$, and the closer is the prior $\Pr[s]$ to 1/2.

The decision maker will take a gamble or is indifferent if

$$E[x] + v(E[\Pr[s|x]) \geq z + v(\Pr[s]) \tag{3}$$

where $v(E[\Pr[s|x]) = \Pr[win]v(\Pr[s|win]) + \Pr[lose]v(\Pr[s|lose])$. If v is strictly concave then by Jensen's inequality the decision maker prefers the prior skill estimate without gambling, $v(\Pr[s]) > v(E[\Pr[s|x])$, making them more wary of gambles than pure monetary considerations would suggest. Analogous to a risk premium, we define the *embarrassment premium* π to be how high a gamble's expected net monetary payoff $E[x] - z$ must be to make the decision maker indifferent to gambling,[8]

$$\pi \equiv v(\Pr[s]) - E[v(\Pr[s|x])]. \tag{4}$$

Hence a "fair gamble" with price $z = E[x]$ will be accepted if $\pi < 0$ and refused if $\pi > 0$. Even for small monetary stakes the fear of looking unskilled remains so this embarrassment premium creates an apparent jump in utility at the status quo if utility is assumed to be a function of wealth alone.

Now consider how the odds of the gamble affect skill updating. Suppose $\Pr[win|s] = \Pr[win] + \Delta/2$ and $\Pr[win|u] = \Pr[win] - \Delta/2$ where the skill gap $\Delta \leq 2\Pr[win], 2\Pr[lose]$ keeps $\Pr[win|s] < 1$ and $\Pr[lose|u] > 0$. For $\Pr[s] = 1/2$ the updated skill estimates from (1) and (2) are $\Pr[s|win] = 1/2 + \Delta/4\Pr[win]$ and $\Pr[s|lose] = 1/2 - \Delta/4\Pr[lose]$. Winning at a long-shot with low $\Pr[win]$ raises estimated skill substantially while losing has only a small impact. Conversely, winning at a sure-thing with high $\Pr[win]$ raises estimated skill only slightly while losing has a large impact.

This differential updating is seen in Figure 1(a) for a long-shot gamble F with $\Pr_F[win] = 1/5$ and sure-thing gamble G with $\Pr_G[win] = 4/5$. Since the gambles have the same degree of outcome uncertainty, $\Pr_F[win]\Pr_F[lose] = \Pr_G[lose]\Pr_G[win]$, it is reasonable to assume the skill gap Δ is the same. We will call two gambles F and G with $\Pr_F[win] = \Pr_G[lose]$ and equal skill gaps "complementary". Given that the skill gap is the same, notice from (1) and (2) that the difference in estimates, $\Pr[s|win] - \Pr[s|lose]$, is inversely proportional to outcome uncertainty $\Pr[win]\Pr[lose]$, so the difference in estimates is the same for the two gambles. The figure shows the particular case where $\Delta = 2\Pr[win]\Pr[lose] = 8/25$ for each gamble.[9]

Losing at a long-shot is not very embarrassing, but happens frequently. And screwing up at a sure-thing is humiliating, but happens rarely. Which is worse? Figure 1(b) shows this tradeoff for the standard constant relative risk aversion function $v = \mu^{1-\eta}/(1-\eta)$ with $\eta = 2$, or $v = -1/\Pr[s|x]$, so $v' > 0$, $v'' < 0$, and $v''' > 0$, implying there is "downside risk aversion" (Menezes, Geiss, and Tressler, 1980), or "prudence" (Kimball, 1990).[10] Since v' is decreasing at a decreasing rate, the sure-thing is worse as seen from the starred expected utilities $E[v(\Pr[s|x])]$ in the figure, implying the embarrassment premium is higher for the sure-thing, $\pi_G = v\left(\frac{1}{2}\right) - \frac{4}{5}v(\frac{3}{5}) - \frac{1}{5}v\left(\frac{1}{10}\right) = \frac{4}{3}$, than for the long-shot, $\pi_F = v\left(\frac{1}{2}\right) - \frac{1}{5}v(\frac{9}{10}) - \frac{4}{5}v\left(\frac{2}{5}\right) = \frac{2}{9}$.

When generally can we rank embarrassment premia for long-shots and sure-things? Figure 2(a) shows CDFs for the respective skill distributions P and Q generated by the long-shot F and sure-thing G from

[8] If U is not linear in y, then the embarrassment premium is just the expected utility from the monetary payoff that makes the decision maker indifferent.

[9] The decreasing lines in the figure show $\Pr[s|w]$ and $\Pr[s|l]$ for different gambles when $\Delta = 2\Pr[w]\Pr[l]$ for every gamble so that the skill gap is the same for gambles with equal uncertainty over the outcome and is larger for gambles with more uncertainty over the outcome. Linear updating in $\Pr[w]$, as shown in the figure, holds for any $\Delta = \alpha\Pr[w]\Pr[l]$ with $\alpha \in (0, 2]$. While not necessary for any of our results, we will show such updating is implicitly assumed by the achievement motivation literature and is consistent with the self-handicapping and prospect theory literatures. More generally, notice from (1) and (2) that $\Pr[s|W]$ and $\Pr[s|L]$ are decreasing in $\Pr[W]$ as long as the skill gap Δ as a function of $\Pr[W]$ does not rise too rapidly, $\Delta'/\Delta < 1/\Pr[W]$.

[10] The conditions $v' > 0$, $v'' < 0$, and $v''' > 0$ are necessary for decreasing absolute risk aversion, which implies that demand for risky assets increases with wealth (Pratt, 1964), imply precautionary savings (Kimball, 1990), and are necessary for the Deusenberry demonstration effect (Harbaugh, 1996).

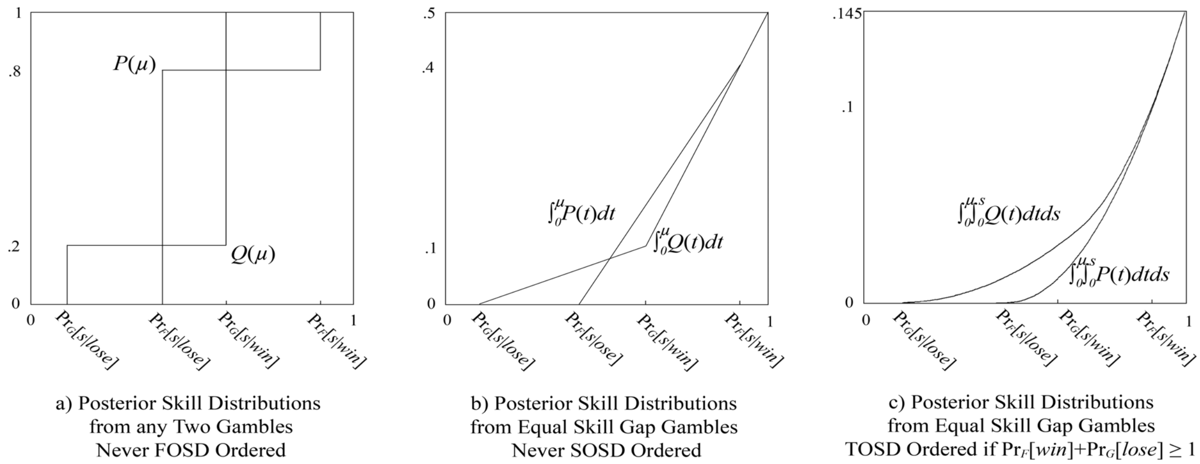


Figure 2: Posterior Skill Distributions Generated by Gambles

the example. The skill distributions from any two gambles cannot be FOSD ranked since the mean estimated skill μ must be the prior $\Pr[s]$, i.e., gambling over skill is a fair gamble. And as seen in Figure 2(b), the skill distributions from any two gambles with an equal skill gap cannot be SOSD ranked because the order of posterior estimates must overlap, $\Pr_G[s|lose] < \Pr_F[s|lose] < \Pr_G[s|win] < \Pr_F[s|win]$ for $\Pr_G[win] > \Pr_F[win]$ from (1).[11]

Regarding Third Order Stochastic Dominance, distribution $P \succ_{TOSD} Q$ if the means are equal (the prior $\mu = \Pr[s]$ in this environment) and the integral of the integral of Q is always higher than that of P, or

$$\int_0^\mu \int_0^s (Q(t) - P(t))\, dtds \geq 0 \tag{5}$$

for all $\mu \in [0, 1]$, implying $\int_0^1 v(\mu)\, dP(\mu) \geq \int_0^1 v(\mu)\, dQ(\mu)$ for $v' > 0$, $v'' \leq 0$, $v''' \geq 0$ (Whitmore, 1970). Looking at Figure 2(c), condition (5) holds in the example so $P \succ_{TOSD} Q$, implying that the long-shot has a lower embarrassment premium than the sure-thing, as calculated above in the example for the functional form $v = -1/\mu$.

When more generally does the skill distribution generated by a long-shot gamble TOSD dominate that generated by a sure-thing?[12] Suppose that, as in the example, we consider complementary gambles F and G where $\Pr_F[win] = \Pr_G[lose]$ and the skill gap is the same. The following lemma uses the diminishing variation property of integration (Karlin, 1967) to show that the long-shot TOSD dominates the sure-thing.

Looking at Figure 2(a), $P(\mu) - Q(\mu)$ crosses zero at most twice. The diminishing variation property implies that the number of zeros does not increase with integration, so the integral $\int_0^\mu (Q(t) - P(t))\, dt$

[11] Note that for gambles with different skill gaps, the gamble with a lower skill gap SOSD dominates if the win probabilities are sufficiently close.

[12] For any two gambles with an equal skill gap the sure-thing gamble with a higher chance of winning can never be TOSD dominant since it has a lower skill estimate from losing, e.g., $\Pr_G[s|L] < \Pr_F[s|L]$ implies $Q(\mu') > P(\mu')$ for some μ', so (5) cannot be positive for all μ' if P and Q are reversed.

shown in Figure 2(b) crosses zero at most twice. The means of each distribution both equal the prior, or $\int_0^1 Q(t)dt = \int_0^1 P(t)dt = \Pr[s]$, and $P(\mu) = Q(\mu)$ for $\mu > \Pr_F[s|win]$, so $\int_0^{\Pr_F[s|win]} (Q(t) - P(t))\, dt = 0$. Therefore, since $Q(\mu) > P(\mu)$ initially, $\int_0^{\mu} (Q(t) - P(t))\, d\mu$ cannot cross zero more than once. This then implies that its integral (5) in Figure 2(c) crosses zero at most once. Since (5) starts positive and is constant above $\mu = \Pr_F[s|win]$, it cannot cross zero at all if (5) is strictly positive at $\mu = \Pr_F[s|win]$ since otherwise it must cross twice, so (5) is always non-negative. As confirmed in the proof, the same is true if (5) equals zero. In either case the value of (5) at $\mu = \Pr_F[s|win]$ can then be checked directly from the areas below $\int_0^{\Pr_F[s|win]} Q(t)dt$ and $\int_0^{\Pr_F[s|win]} P(t)dt$ in Figure 2(b) as done in the proof, with the following result.

Lemma 1 *For gambles F and G with equal skill gaps $\Delta_F = \Delta_G$, the posterior skill distributions Q and P satisfy $P \succ_{TOSD} Q$ if $\Pr_F[win] < \Pr_G[win]$ and $\Pr_F[win] + \Pr_G[win] \geq 1$.*

For a long-shot F and complementary sure-thing G where $\Pr_F[win] = \Pr_G[lose]$, the condition $\Pr_F[win] + \Pr_G[win] \geq 1$ is satisfied with equality, so TOSD holds. Therefore $E_F[v(s|x)] > E_G[v(s|x)]$ under the downside risk aversion assumptions on v, implying from (4) that the long-shot has a lower embarrassment premium. Note that strict satisfaction of the condition will be relevant for applying the lemma in subsequent sections where the decision maker has private information and is more likely to gamble when this information is favorable.[13]

We summarize these results as follows.

Proposition 1 *For performance skill without private information: (i) the embarrassment premium π is always positive, and (ii) the embarrassment premium π is lower for long-shot gamble F than complementary sure-thing gamble G.*

A higher risk premium for sure-things implies that, for equal prices z, a decision maker who takes a sure-thing will always take the complementary long-shot, but not vice versa. So gambling will be observed for a wider range of prices for a long-shot than sure-thing.

1.2 Performance Skill: Decision Maker has Private Information

Real world gambles have many complexities absent from this simple model. Of particular interest for connecting our results to the social psychology and behavioral literatures, the decision maker might have some indication of their own skill, so refusing to take a gamble can itself be informative. We therefore generalize the above approach to allow the decision maker to have a private signal $\theta \in \{g, b\}$ of their skill, $\Pr[s|g] > \Pr[s|b]$. Since skill is correlated with winning, $\Pr[win|s] > \Pr[win|u]$, the

[13] When $\Pr_F[W] + \Pr_G[W] < 1$, e.g., both gambles are long-shots, gamble G is closer to an even bet than gamble F, which implies from (1) and (2) that the gap between skill estimates from winning and losing is higher for F. This adds to the riskiness of the gamble, counteracting the effect of less extreme embarrassment from losing at F as captured by $\Pr_F[s|L] > \Pr_G[s|L]$. In particular for equal skill gaps $(\Pr_F[s|W] - \Pr_F[s|L]) - (\Pr_G[s|W] - \Pr_G[s|L])$ is proportional to $\Pr_G[W]\Pr_G[L] - \Pr_F[W]\Pr_F[L]$, which for $\Pr_F[W] < \Pr_G[W]$ is positive if $\Pr_F[W] + \Pr_G[W] < 1$ and negative if $\Pr_F[W] + \Pr_G[W] > 1$.

signal is also, $\Pr[win|g] > \Pr[win|b]$.[14] Conditional on skill a, θ provides no information on winning, $\Pr[win|a, \theta] = \Pr[win|a]$. The gamble is now defined by its distribution $F(x, a, \theta)$ which has full support.

The decision maker is now a sender in a signaling game where the choice to gamble or not can signal the sender's type θ to a receiver. Our equilibrium concept is sequential equilibrium where F and the opportunity cost or price z are common knowledge, and where observer beliefs about which types gamble and which types refuse are consistent with decision maker strategies on the equilibrium path and equal the limiting beliefs of a mixed strategy equilibrium off the equilibrium path. In the text we will focus on the separating equilibrium and analyze pooling and mixed strategy equilibria, with additional notation for strategies and beliefs, in the Appendix.

Since the private signal θ is informative of skill, the tradeoff from gambling for each type will differ, and there can be different equilibria with different equilibrium beliefs about which type(s) gamble. If a gamble is believed to be type θ then, generalizing (1) and (2), updated skill is

$$\Pr[s|x, \theta] = \Pr[s|\theta] + \frac{\Pr[x|s] - \Pr[x|u]}{\Pr[x|\theta]} \Pr[s|\theta] \Pr[u|\theta]. \tag{6}$$

A separating equilibrium where only type g gambles exists if the payoffs given such beliefs from gambling $E[x|\theta] + E[v(\Pr[s|x, g])|\theta]$ and not gambling $z + v(\Pr[s|b])$ make g prefer gambling and b prefer not gambling. Rearranging, and generalizing the embarrassment premium introduced above to allow for the signal θ, the equilibrium exists if the embarrassment premium is lower than any monetary gain for g but not for b,

$$\pi_g \equiv v(\Pr[s|b]) - E[v(\Pr[s|x, g])|g] \leq E[x|g] - z \text{ and} \tag{7}$$

$$\pi_b \equiv v(\Pr[s|b]) - E[v(\Pr[s|x, g])|b] \geq E[x|b] - z. \tag{8}$$

Notice that not gambling indicates a bad signal b and induces an unfavorable but certain skill estimate $\Pr[s|b]$, while gambling indicates a good signal g and induces a more favorable but also more variable estimate $\Pr[s|x, g]$. Since $\Pr[win|g] > \Pr[win|b]$, both the skill and monetary components are better for g, $E[v(\Pr[s|x, g])|g] > E[v(\Pr[s|x, g])|b]$ and $E[x|g] > E[x|b]$, so a separating equilibrium exists for some z. For this same reason each type can be made indifferent with a different z while still in the separating equilibrium.

The embarrassment premium for each type of decision maker shows the tradeoff between admitting incompetence by not gambling and risking embarrassment by gambling. The tradeoff favors not gambling when the private signal θ indicates little or nothing about skill as in the above section – as $\Pr[s|g] - \Pr[s|b]$ approaches 0, the embarrassment premium π_θ for each type approaches (4) and hence is positive. Conversely, the tradeoff instead favors gambling when θ is very informative about skill so refusing is very revealing of low skill. As $\Pr[s|g] - \Pr[s|b]$ goes to 1, $\Pr[s|x, g]$ goes to 1 and $\Pr[s|x, b]$ goes to 0, so the premium goes to

$$\pi_\theta = v(0) - v(1) < 0. \tag{9}$$

[14] Since the "signal" to gamble or not is naturally binary, we keep the other variables binary as well. The assumptions of binary skill a and a binary outcome x, which are consistent with most of the related literature, can be relaxed. However allowing for a richer type space for θ when the signal is binary introduces extra complications (e.g., Adriani and Sonderegger, 2019).

The tradeoff also favors gambling when the skill gap is sufficiently small. As $\Pr[win|s] - \Pr[win|u]$ goes to 0 and $\Pr[s|x]$ goes to $\Pr[s]$, the premium goes to

$$\pi_\theta = v(\Pr[s|b]) - v(\Pr[s|g]) < 0. \tag{10}$$

In this case there is little loss in estimated skill from losing, so taking a fair gamble is less embarrassing than admitting to a lack of confidence by refusing.[15] This establishes parts (i) and (ii) of Proposition 2 below for the separating equilibrium case.

Still focusing on the separating equilibrium, we are again interested in comparing embarrassment premia for a long-shot F and sure-thing G where the gambles are complementary in that $\Pr_F[win] = \Pr_G[lose]$ and $\Delta = \Pr_F[win|s] - \Pr_F[win|u] = \Pr_G[win|s] - \Pr_G[win|u]$ as before, plus the signal information remains the same across gambles $\Pr_F[s|\theta] = \Pr_G[s|\theta]$. Looking again at Lemma 1, notice that we can condition F and G on $\theta = g$ and the lemma applies directly where P is the posterior distribution from gamble $F(x, a|g)$ and Q is the posterior distribution from gamble $G(x, a|g)$. For the type g decision maker all of the analysis proceeds as above but the probabilities are conditional on $\theta = g$ as in (6) and (7). So the long-shot has a lower risk premium, $\pi_{F,g} < \pi_{G,g}$, if $E[v(\Pr_F[s|x, g])|g] > E[v(\Pr_G[s|x, g])|g]$, which holds by the above analysis if $\Pr_F[win|g] + \Pr_G[win|g] \geq 1$. Since $\Pr[win|g] > \Pr[win]$ for both gambles, this condition holds if $\Pr_F[win] + \Pr_G[win] = 1$, as assumed for complementary gambles. For type b, there is higher weight on both $v(\Pr_F[s|lose, g])$ and $v(\Pr_G[s|lose, g])$ in (8) than there is for type g in (7). Since the latter is more negative from (6), the long-shot also has a lower risk premium.

Regarding other equilibria, if z is sufficiently low a both-gamble equilibrium exists and if z is sufficiently high a both-refuse equilibrium exists, and for some intermediate values partial pooling equilibria exist. Unlike the separating equilibrium, in these other equilibria an embarrassment premium can be measured for only one type. In the both-gamble equilibrium indifference by the b type implies the g type cannot be made indifferent so only π_b can be measured, while in the neither-gamble equilibrium indifference by the g type implies the b type cannot be made indifferent so only π_g can be measured. Similarly, for partial pooling equilibria π_θ is measurable only for the indifferent pooling type θ. The proof in the Appendix extends the separating equilibrium results to show that the Proposition below holds for any π_θ that is measurable in an equilibrium.

Proposition 2 *For performance skill with private skill signal θ, in any equilibrium, the measurable embarrassment premia π_θ are (i) positive for sufficiently weak signal θ, (ii) negative for sufficiently strong signal θ or sufficiently small skill gap Δ, and (iii) lower for long-shot F than complementary sure-thing G.*

Since long-shots offer a lower embarrassment premium in any equilibrium, for any given price z if a type is willing to gamble on a sure-thing it would also be willing to gamble on a complementary long-shot. And, if a type is unwilling to gamble on the long-shot it would also be unwilling to gamble on the sure-thing at the same price. Hence, even though multiple equilibria exist, this provides the testable prediction that the range of prices z that supports gambling is larger for a long-shot than a sure-thing.

[15] See Chen (2016) for a more general analysis of the signaling incentive to take a risky project. Chung and Eso (2013) analyze the incentives to simultaneously show off and also learn about one's ability.

1.3 Evaluation Skill: Outcome of Gamble Observed Only if Accepted

Instead of being more likely to succeed at a given project, a skilled decision maker may be more likely to identify projects with better odds of success.[16] Success at a gamble can then indicate that the decision maker had a more accurate estimate of the odds, and hence is more skilled. With performance skill the outcome of a refused gamble does not reflect on the decision maker's skill. But with evaluation skill success or failure of a gamble not taken can still reflect on their judgement. The value of an unpurchased asset still rises or falls even if it is not bought, so forgoing an opportunity that does well can be as embarrassing as making an investment that goes poorly. Hence for evaluation skill we consider both the case where the outcome of a refused gamble is not observed and the case where it is observed.

We first consider the case where only the outcomes of an accepted gamble are observed, and again we focus in the text on the separating equilibrium where only g gambles. With evaluation skill the signal is informative of winning, $\Pr[win|g] > \Pr[win|b]$, but now θ provides no direct information on skill, $\Pr[s|g] = \Pr[s|b]$, and skill unconditional on a signal is independent of winning, $\Pr[win|a] = \Pr[win]$. Skill matters because the decision maker's signal θ is more informative for a skilled decision maker, so the skill gap for the separating equilibrium where g gambles and b does not is $\Delta \equiv \Pr[win|s,g] - \Pr[win|u,g] = \Pr[lose|s,b] - \Pr[lose|u,b] > 0$ and estimated skill from (6) is now

$$\Pr[s|x,\theta] = \Pr[s] + \frac{\Pr[x|s,\theta] - \Pr[x|u,\theta]}{\Pr[x|\theta]} \Pr[s]\Pr[u]. \tag{11}$$

For instance, consider the limiting case where a skilled decision maker's signal θ is always accurate of the true probability and an unskilled decision maker's signal is noise. So letting suppose the true probability of winning is equally likely to be either $\Pr[win]+\Delta$ or $\Pr[win]-\Delta$. Then, since only g types are expected to gamble, $\Pr[win|s,g] = \Pr[win]+\Delta$ and $\Pr[win|u,g] = \Pr[win]$, so $\Delta > 0$ is the skill gap. Since $\Pr[win|g] = (\Pr[win|s,g]\Pr[s,g] + \Pr[win|u,g]\Pr[u,g]) / \Pr[g] = \Pr[win]+\Pr[s]\Delta$, estimated skill is $\Pr[s|win,g] = \Pr[s] + (\Delta/(\Pr[win]+\Pr[s]\Delta))\Pr[s]\Pr[u]$. Figure 3(a) show this case when $\Delta = \Pr[win]\Pr[lose]$ and $\Pr[s] = 1/2$, so $\Pr[s|win,g] = 1/2 + 4\Pr[win]\Pr[lose]/(\Pr[win]+\Pr[win]\Pr[lose]/2)$ and, similarly, $\Pr[s|lose,g] = 1/2 - 4\Pr[win]\Pr[lose]/(\Pr[win]+\Pr[win]\Pr[lose]/2)$.

The observer learns nothing from the decision to not gamble since the decision maker's signal θ contains no direct information about ability. So expected skill from not gambling is just the prior $\Pr[s|g] = \Pr[s|b] = \Pr[s]$ in any equilibrium. The separating equilibrium where g gambles exists if

$$\pi_g \equiv v(\Pr[s]) - E[v(\Pr[s|x,g])|g] \leq E[x|g] - z \text{ and} \tag{12}$$

$$\pi_b \equiv v(\Pr[s]) - E[v(\Pr[s|x,g])|b] \geq E[x|b] - z. \tag{13}$$

For type g the expected skill estimate from gambling is $\Pr[s]$ but the realized estimate is uncertain so the embarrassment premium is positive, $\pi_g > 0$. For type b the expected skill estimate is even worse since $\Pr[win|b] < \Pr[win|g]$ so the premium is also positive, $\pi_b > 0$.

To compare embarrassment premia for long-shots and sure-things, again consider the separating equilibrium where the observer expects only type g to gamble and only that type does gamble. By the same argument as above we can apply Lemma 1 by conditioning on $\theta = g$ so P is the posterior distribution

[16] In the career concerns literature such evaluation skill has been used to understand problems ranging from distorted investment decisions (Holmstrom, 1982) to political correctness (Stephen Morris, 2001).

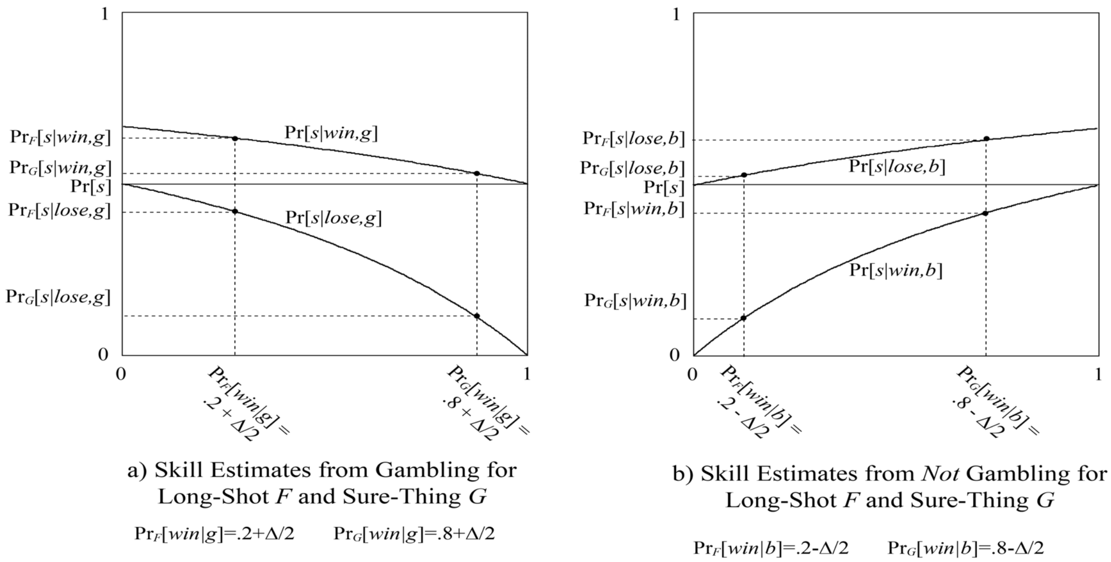


Figure 3: Evaluation Skill: Posterior Skill Estimates

from gamble $F(x, a|g)$ and Q is the posterior distribution from gamble $G(x, a|g)$. Then based on (11) and (12), $\pi_{F,g} < \pi_{G,g}$ if $E[v(\Pr_F[s|x, g])|g] > E[v(\Pr_G[s|x, g])|g]$, which holds by the above analysis if $\Pr_F[win|g] + \Pr_G[win|g] \geq 1$. Since $\Pr[win|g] > \Pr[win]$ for both gambles, this condition holds if $\Pr_F[win] + \Pr_G[win] = 1$ as is true if the gambles are complementary. As in the performance skill case, the result then extends to type b as confirmed in the proof.

With evaluation skill a skilled decision maker is equally good at recognizing a bad gamble, so betting against success is an option. We do not allow the decision maker to choose which side of the outcome to bet on financially, e.g., to place a put option, but there can be equilibria where there is a financial loss from losing but a flipped reputational gain from losing that encourages b types to gamble. If the receiver believes a gamble was primarily from b types whose signal suggests that losing is more likely than expected, then losing is a signal of skill, which then gives b types more incentive to gamble that can confirm the receiver's beliefs in equilibrium. In these equilibria the skill gap is endogenously reversed as shown in the Appendix.

While such equilibria are interesting we focus on "standard" equilibria in which gambling is believed to be at least as likely from g types than b types. The following extends the results from the above separating equilibrium to all such standard equilibria.

Proposition 3 *With evaluation skill when the outcome of a refused gamble is not observed: (i) the embarrassment premia π_θ in all standard equilibria are non-negative (ii) the embarrassment premia π_θ are lower in any standard equilibrium for long-shot F than complementary sure-thing G.*

These risk premia results are consistent with the insight from the performance skill results that we expect to see more gambling with long-shots than sure-things. Such a pattern holds as long as we focus on standard equilibria, but otherwise such a pattern is not ensured.

1.4 Evaluation Skill: Outcome of Refused Gamble Also Observed

When a missed opportunity goes well, the decision maker looks bad. So if the outcomes of both accepted and refused gambles are observable, it is not only risky to take a gamble, but also risky to refuse it. Which is worse? Figure 3(b) shows the same separating equilibrium example as in Figure 3(a) except the skill estimates are now the updated estimates conditional on not gambling rather than gambling. Comparing the two panels, notice that for long-shots it is more embarrassing to not take the gamble and see it win than to take the gamble and see it lose, and the opposite for sure-things.

Given that not gambling at a long-shot can be riskier than gambling, is it better to gamble? As seen in Figure 3, and as we verify more generally in the proof of Proposition 4 below, in a separating equilibrium where only g gambles at given gamble F the skill distribution for type g from deviating and not taking gamble F equals the skill distribution for type b from taking a complementary gamble G. Similarly the skill distribution for type b from not taking gamble F equals the skill distribution for type g from taking a complementary gamble G. Therefore a separating equilibrium where g gambles and b refuses exists for gamble F if

$$\pi_g \equiv E[v(\Pr_G[s|x,g])|b] - E[v(\Pr_F[s|x,g])|g] \leq E_F[x|g] - z \text{ and} \tag{14}$$

$$\pi_b \equiv E[v(\Pr_G[s|x,g])|g] - E[v(\Pr_F[s|x,g])|b] \geq E_F[x|b] - z. \tag{15}$$

So in evaluating the choice to take a gamble F or not, the skill estimate component of the choice is equivalent to evaluating whether F or its complementary gamble G has a lower embarrassment premium. In particular the average embarrassment premium for types g and b is the sum of (14) and (15), or $\pi_F = E[v(\Pr_G[s|x,g])] - E[v(\Pr_F[s|x,g])]$. Noting that $\Pr_F[win|x,g] + \Pr_G[win|x,g] > 1$, Proposition 3(ii) implies $E[v(\Pr_G[s|x,g])|\theta] - E[v(\Pr_F[s|x,g])|\theta] < 0$ if F is a long-shot and the opposite if F is a sure-thing, so π_F is negative or positive accordingly.[17] This establishes Proposition 4(i-a) below, which is stronger than previous results in that long-shots are not just favored over sure-things, but have negative versus positive risk premia.

Now consider pooling equilibria. In the both-gamble equilibrium, winning or losing at the gamble reveals nothing about the decision maker's signal and hence whether the signal was right or wrong, so expected skill stays at $\Pr[s]$. If the decision maker deviates and refuses, then for any belief about who deviated that differs from the prior $\Pr[g]$, the signal combined with the outcome provides information on expected skill. For instance if the deviation is believed to come from type b, failure is a good sign of skill and success a bad sign, so $\Pr[s|x]$ will vary from its mean of $\Pr[s]$, implying by $v'' < 0$ that the embarrassment premium is negative for all $\Pr[win]$. Similarly in the both-refuse equilibrium, if the refused gamble wins or loses, expected skill stays the same but a deviation to gambling is risky. Hence the embarrassment premium is instead positive for all $\Pr[win]$.[18]

Proposition 4 *With evaluation skill when the outcome of a refused gamble is observed: (i) in the separating equilibrium where g types gamble and b types refuse the average embarrassment premium π*

[17] This is for the unweighted average premium. For the average weighted by $\Pr[g]$ and $\Pr[b]$ to flip exactly at $\Pr[w] = 1/2$ requires $\Pr[g] = \Pr[b]$.

[18] For any partial pooling equilibria the embarrassment premia are ambiguous since they reflect opposing factors from the separating and pooling equilibria. For sufficiently small monetary stakes, there can also be non-standard equilibria as discussed in the previous section.

is negative for $\Pr[win|g] < 1/2$ *and positive for* $\Pr[win|b] > 1/2$, *(ii) in the both-gamble equilibrium the embarrassment premium* π_θ *is negative or zero, and (iii) in the both-refuse equilibrium the embarrassment premium* π_θ *is positive or zero.*

These results are consistent with the favoring of long-shots in the previous, but also highlight the strong sensitivity of behavior to observer beliefs.

2 Rationalization of Behavioral Models

An aversion to losing appears in a wide variety of behavioral models. A model of Bayesian updating based on the career concerns and social signaling approaches can rationalize many of the general insights and particular functional forms used in these models.

2.1 Self-Esteem

The idea that self-esteem depends on the outcomes of risky decisions, and that people may avoid risk to protect their self-esteem, dates back at least to James (1890) who defined self-esteem as the ratio of successes to "pretensions". He noted that self-esteem could be raised both by "increasing the numerator" through success and by "diminishing the denominator" through avoidance. A preference for greater self-esteem corresponds to $v' > 0$, while James' suggestion that protecting self-esteem drives behavior corresponds to $v'' < 0$.[19] Goffman (1959) analyzes strategies for managing the esteem of others, and highlights that embarrassment is driven by unexpected failure.[20] If self-esteem does not depend on observer inferences, then the basic insights of the introductory example still apply, but the asymmetric information analysis in Section 3 would require some form of self-signaling.[21]

2.2 Achievement Motivation

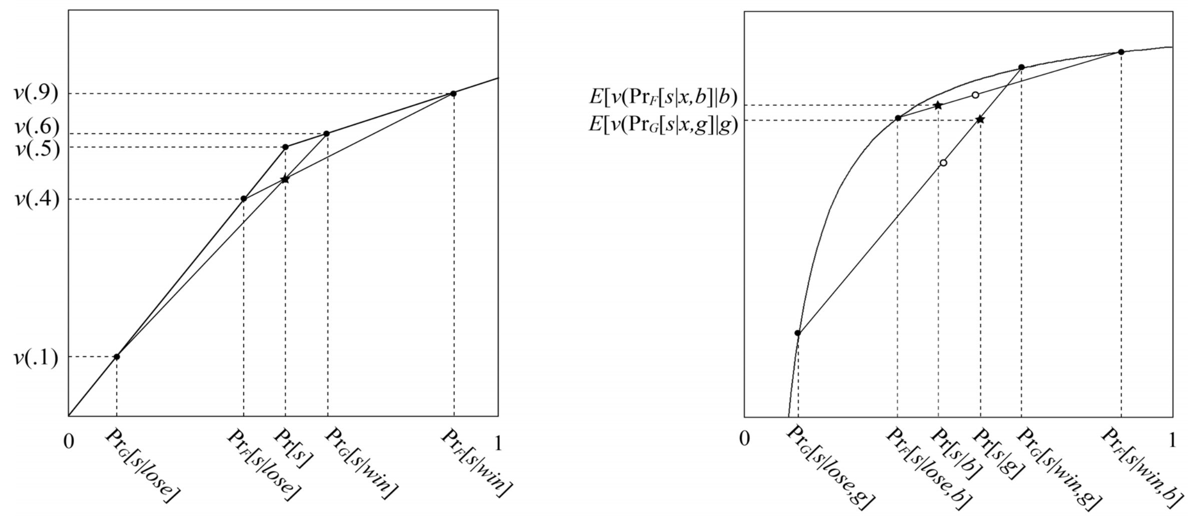
A leading theory of risk taking before prospect theory, Atkinson's (1957) theory of achievement motivation captures the idea that different probability gambles convey different information about skill. Consistent with the model, experiments with an explicit skill component found that people were afraid of gambles with an equal probability of success or failure. However, experiments also found a strong tendency to favor long-shots over sure-things, which was considered to be outside of the model's predictions (e.g., Atkinson et al., 1960).

Atkinson (1957) assumes that, for a gamble with chance p chance of success, the utility from success is $m_s(1-p)$ and the utility from failure is $m_f(-p)$ where the constants $m_f > m_s > 0$ reflect the respective motives to avoid failure and gain success. Noting that the utility gain from winning is higher when p is

[19] Cowen and Glazer (2006) consider labor market applications where risk aversion with respect to ability estimates is likely.

[20] Self-esteem can be instrumental if it facilitates conveying a favorable image to others (Benabou and Tirole, 2002). Burks et al. (2013) find evidence consistent with over-confidence as a social signal.

[21] Self-signaling in this context can be based on intrapersonal asymmetric information (Benabou and Tirole, 2002 and 2004; Bodner and Prelec, 2003).



a) Achievement Motivation without Downside Risk Aversion

b) Self-Handicapping with Private Information

Figure 4: Formalization of Early Embarrassment Aversion Models

small, while the utility loss from losing is higher when p is large, the expected utility from the gamble is

$$pm_s(1-p) + (1-p)m_f(-p), \tag{16}$$

implying the utility from gambling is lowest at $p = 1/2$.

From the perspective of embarrassment aversion, that the utility from winning increases in the probability of winning p while the disutility from losing increases in p is consistent with a concern for estimated skill based on Bayesian updating. In particular such a model arises when $\Delta = \alpha p(1-p)$ so skill estimates are linear in the probability of success as in the introductory example and v is a piecewise linear function with slope m_s above the prior $\Pr[s]$ and $-m_f$ below. From (1) and (2), and normalizing $v(\Pr[s]) = 0$, $v = m_s \Pr[s|win]$ for $x = win$ and $v = -m_f \Pr[s|lose]$ for $x = lose$, expected v is then

$$\begin{aligned} v &= pm_s \frac{\Delta}{p} \Pr[s]\Pr[u] - (1-p)m_f \frac{\Delta}{(1-p)} \Pr[s]\Pr[u] \\ &= pm_s(1-p)\alpha \Pr[s]\Pr[u] + (1-p)m_f(-p\alpha \Pr[s]\Pr[u]), \end{aligned} \tag{17}$$

where the only difference from (16) is that the motives m_f and m_s are amplified by a larger skill gap (higher α) and dampened by a stronger prior (high $\Pr[s]$ or $\Pr[u]$).

As seen in Figure 3(a) where $m_f = 4/3$, $m_s = 1/6$ and $\Pr[s] = 1/4$ the linear segments make v concave and hence consistent with loss aversion, but preclude a role for downside risk aversion. If m_f were increasing and m_s decreasing rather than constant the model would allow for $v''' > 0$, which from Proposition 2 would then fit the experimental pattern of favoring long-shots over sure-things observed in the early achievement motivation experiments without monetary payoffs.

2.3 Self-Handicapping

A preference for long-shots is central to the question of self-handicapping. To reduce the loss in esteem due to failure, people deliberately lower the odds of success so that failure is likely (Jones and Berglas, 1978). The literature considers both self-esteem and esteem by others as factors and finds that self-handicapping is more common in public situations (Kolditz and Arkin, 1982). Our model with performance skill formalizes implicit assumptions in the literature. First, losing at lower probability gambles is indeed less damaging to estimated skill as shown in the introductory example. Second, this gain can compensate for more frequent loss if there is downside risk aversion.[22] Third, even if the choice to self-handicap is itself an embarrassing signal, we find that self-handicapping can still be an equilibrium.

Exploring this last issue, in the initial example of performance skill without an informative signal the embarrassment premium is lower for long-shots, so the decision maker is better off self-handicapping if the monetary loss is smaller than the expected utility difference between long-shots and sure-things shown in Figure 1(b). Hence, applying the logic of Proposition 1(i) and Proposition 2, both types will still want to self-handicap if the signal θ is weak and the stakes are small. Since they both do so, self-handicapping is less embarrassing than if only the b type does.

To see how a separating equilibrium can arise where only type b self-handicaps from a sure-thing G to a complementary long-shot F, note that the choice in this context is not whether to take a gamble but which of the two gambles to take. Letting the prices z and material stakes x of the gambles be the same, note that $E_G[x|g] > E_G[x|b]$ and $E_F[x|g] > E_F[x|b]$ since $\Pr[win|g] > \Pr[win|b]$, but the expected gain from a good signal g is the same for each gamble so $E_G[x|g]-E_F[x|g] = E_G[x|b]-E_F[x|b] = y' > 0$. That is they both face the same expected material gain y' from sticking with G and not self-handicapping. Because of the curvature of v due to embarrassment aversion, the expected utility from losing at G is relatively worse for b than g. Therefore, for some y' it is possible to satisfy the separating equilibrium conditions

$$y' \geq E_F[v(\Pr_F[s|x,b])|g] - E_G[v(\Pr_G[s|x,g])|g] \text{ and} \tag{18}$$

$$y' \leq E_F[v(\Pr_F[s|x,b])|b] - E_G[v(\Pr_G[s|x,b])|b]. \tag{19}$$

Figure 4(b) shows the introductory example of Section 2.1 of $\Pr_F[win] = 1/5$ and $\Pr_G[win] = 4/5$, but with private information as in as in Section 2.2, $\Pr[s|g] = 5/9$ and $\Pr[s|b] = 4/9$. In the separating equilibrium taking the original gamble G is a good signal that raises estimated skill, while self-handicapping down to gamble F is a bad signal that lowers estimated skill. The left star and circle represent the equilibrium and deviation payoffs for b, and the right pair for g. For both types gamble F is preferred based on the expected utility from skill estimates, but b gains more than g does from self-handicapping, so if y' is between the two differences the equilibrium holds.[23]

Indeed it is possible for a separating equilibrium to exist even when type g is not hurt materially from self-handicapping, $y' = 0$. As the signal θ becomes more accurate, a b type is more and more likely to be

[22] Benabou and Tirole (2002) analyze self-handicapping as an inefficient action that completely avoids revealing ability rather than reducing the probability of success.

[23] Note though that gamble G offer a slightly higher average skill since b would be pooling with g types. This is seen from the circled deviation payoff for type b taking gamble G being a convex combination of the winning and losing skill estimates with an average estimate slighty above $\Pr[s|b]$. Similarly if type g deviates to gamble F, then the average estimate is slightly below $\Pr[s|g]$. With a stronger signal θ these differences increase as discussed below.

unskilled. So a g type does better by showing off and sticking with the long-shot G even at the chance of embarrassment than by taking the easier F gamble and appearing to admit a lack of confidence. From Figure 4(b), the equilibrium star point for the g type moves to the right as their g signal becomes more accurate, but the deviation circle payoff moves to the left as the choice to self-handicap becomes a stronger signal of a lack of skill. Type b would also like to show confidence, but the chance of an embarrassing failure is high given the b signal, so if the signal θ is not too accurate they are better off admitting to their unfavorable signal b and self-handicapping than by proving their lack of skill by failing at the sure-thing.[24]

2.4 Prospect Theory

Prospect theory incorporates a range of behavioral anomalies relating to risk-taking. We focus on how embarrassment aversion relates to loss aversion, probability weighting, and framing.[25]

Loss Aversion: Researchers have long noted that people are risk averse even for very small gambles where a smooth utility function predicts local risk neutrality. Prospect theory captures this phenomenon by assuming that utility from monetary outcomes is kinked at the status quo (or other reference point) so the marginal pain from losing is strictly greater than the marginal gain from winning even as the stakes go to zero.

If decision makers are embarrassment averse with $U = y + v(\mu)$, or more generally if $U = u(y) + v(\mu)$ for smooth function u, then even as the monetary stakes x go to zero, and hence variation in y goes to zero, utility from winning or losing is still affected by estimated skill μ. For no private information about skill by Proposition 1, or for limited private information by Proposition 2, there will still be positive embarrassment premia for small-stake gambles even for (locally) linear monetary utility function u. Looking at Figure 4(a), if the gamble at price z is not taken then normalize $y = z$ so $U = z + v\,(\Pr[s])$. If the gamble is won then income increases by win, and utility moves up by win also shifts upward by $v(\Pr[s|win]) - v\,(\Pr[s])$, while if it is lost then utility moves down by $lose$ and also shifts down to by $v(\Pr[s]) - v(\Pr[s|lose])$. Hence even though U is smooth in y and estimated skill μ, if v is not included in total utility then U will appear to be kinked, or in fact discontinuous as the gamble stakes go to zero, at the status quo. As seen in the figure, concavity of v implies $v\,(\Pr[s]) > E\,[v\,[\Pr[s|x]]$ so not gambling offers higher expected utility even at the limit as the stakes equal $win = lose = 0$.

With prospect theory as the stakes go to zero there is still risk aversion due to the kink, but the actual risk premium goes to zero, and at the limit the decision maker is indifferent to gambling or not. But with embarrassment aversion the premium does not disappear since $v\,(\mu)$ remains a factor – decision makers still dislike appearing unskilled even at a "friendly bet" with no monetary stakes. Hence a kink can roughly capture positive risk premia, but for arbitrarily small gambles a discontinuity is necessary as seen in Figure 4.

[24] In the above example the g type prefers G for accuracy $\Pr[s|g] = 1 - \Pr[u|g]$ greater than .57, and b also prefers G for accuracy greater than .59, so a separating equilibrium exists in the intermediate range. This gap increases as the difference between the win probabilities for the sure-thing and long-shot increases.

[25] The results are also related to regret theory (Bell, 1982; Loomes and Sugden, 1982), rank-dependent utility (Quiggin, 1982), and disappointment aversion (Gul, 1991) via their known connections to prospect theory. Bell (1982) notes "the evaluation of others, one's bosses for example, may be an important consideration" in regret. Steiner and Stewart (2016) analyze probability weighting as a rational correction for the winner's curse.

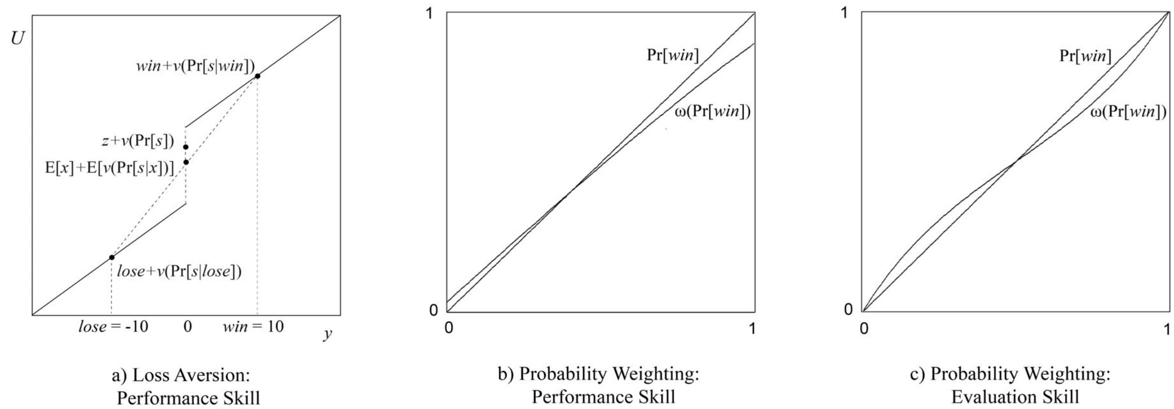


Figure 5: Embarrassment Aversion Mapped to Prospect Theory's Probability Weights

Framing: Prospect theory's kinked value function is assumed to generate risk aversion in the region above the status quo or reference point (positive domain) but risk loving behavior in the lower region below the reference point (negative domain). Whether outcomes are perceived to be in one region or the other can then be manipulated by framing of the gamble. For a fair gamble, when losing is framed as the reference point, not gambling is a gain, so risk aversion results. But when winning is framed as the reference point, not gambling is a loss, so risk loving results.[26]

Embarrassment aversion allows framing to similarly affect behavior by affecting equilibrium expectations and hence "selecting" among the multiplicity of equilibria in signaling games. For performance skill, gambling can be presented as an admission of low skill, which is consistent with a separating equilibrium or both-gamble equilibrium, or be presented neutrally, which is consistent with a neither-gamble equilibrium. For evaluation skill either pooling equilibrium hides any evidence of skill, so as seen in Proposition 4 the decision maker has an incentive to follow whatever they think is expected rather than risk an embarrassing failure.[27]

Probability Weighting: Prospect theory assumes that people overweight small probability gains (pay \$10 for a 10% "long-shot" chance to win \$100) and underweight high probability losses (risking a 90% chance of losing \$100 to "win back" money rather than pay \$90 for sure), while they also underweight high probability gains (taking \$90 over a 90% "sure-thing" chance of winning \$100) and overweight low probability losses (pay \$10 "insurance" rather than risk a 10% chance of losing \$100) (Kahneman and Tversky, 1979).[28] From the perspective of embarrassment aversion, the first two cases reduce to overweighting a low probability of success, and the last two to underweighting a high probability of success.

The probability weighting function is typically estimated by finding the certainty equivalent z^* that

[26] The four-fold pattern of the probability weighting function makes a similar prediction. Gonzalez and Wu (xx) distinguish the exact predictions effects from those generated by the loss function.

[27] In the classic flu problem (Tversky and Kahneman, 1981) a description (and potential newspaper headline) of "people will be saved" versus "people will die" suggests different expectations even if the number of deaths is the same.

[28] This pattern is sometime referred to the "four-fold pattern of probability", where the differences in the gains and losses domains are attributed to the reflection effect of the value function.

induces indifference to the gamble for different win probabilities $\Pr[win]$ and then inferring what weighted probability $\omega(\Pr[win])$ would induce indifference by a risk neutral decision maker, $\omega(\Pr[win]) \cdot win + (1 - \omega(\Pr[win])) \cdot lose = z^*(\Pr[win])$, implying $\omega(\Pr[win]) = (z^*(\Pr[win]) - lose) / (win - lose)$. In our model $z^* = \Pr[win] \cdot win + (1 - \Pr[win]) \cdot lose - \pi$ where π is the embarrassment premium that depends on $\Pr[win]$. If $\omega(\Pr[win])$ is estimated based on assuming $U = y$ while the true utility function is $U = y + v(\mu)$ without probability weighting,[29] then

$$\omega(\Pr[win]) = \Pr[win] - \frac{\pi(\Pr[win])}{win - lose}, \tag{20}$$

so there appears to be underweighting or overweighting depending on the sign of the embarrassment premium, and such weighting is moderated by higher monetary stakes. There is always underweighting if the private signal is sufficiently weak as in the introductory example by Proposition 1(i), but less underweighting for long-shots by Proposition 1(ii).

With a stronger private signal of ability, Proposition 2(ii) implies it is better to take a chance and gamble than admit incompetence if the skill gap is sufficiently small, implying a negative embarrassment premium and hence probability overweighting.[30] Considering the separating equilibrium, assume the same parameters as in Figure 1 except, as in the self-handicapping example, $\Pr[s|g] - \Pr[s|b] = 1/10$, and set $win = 10, lose = 0$. As seen in Figure 4(b), now there is overweighting of low probability gambles as in the canonical form of Kahneman and Tversky (1979).[31]

For evaluation skill, the embarrassment premium is lower for long-shots than sure-things by Proposition 3(ii), implying relative overweighting for long-shots. And if the outcomes of refused gambles are observed there is overweighting of long-shots and underweighting of sure-things by Proposition 4(i). Considering this latter case, and assuming the same parameters for the example in Section 2.2, the imputed probability weighting function for the example of Figure 4(c) with $win = 1, lose = 0$ is similar to that in Tversky and Kahneman (1992).[32]

3 Conclusion

Economic models based on immediate monetary payoffs are often poor predictors of behavior. One approach to this failing is to maintain the focus on these payoffs but incorporate perceptual and cognitive biases that interfere with rational decision making. Another approach, followed by the early social psychology literature and the more recent information economics literature, is to add to the model information effects that are often of practical importance. The information economics literature has shown how a wide range of seemingly irrational behavior can arise from the interactions of self-interested individuals with different information. Following this general literature, and the career concerns literature

[29] Probability weights in prospect theory are usually measure by assuming linearity of utility (or value) in wealth. Wu and Gonzalez (1996) disentangle the predictions of the probability weighting function and the convex-concave utility function assumed in original prospect theory.

[30] For stronger signals, for both performance and evaluation skill, π can be large enough for the "uncertainty effect" of $\omega(p) < 0$ found by Gneezy, List, and Wu (2006).

[31] In this example there is a "certainty effect" (Kahneman and Tversky, 1979) or distinct behavior at $p = 1$, which arises with embarrassment aversion if the skill gap declines sufficiently slowly as p goes to 1 (see Wakker, 2010, p. 210).

[32] The exact pattern depends on the parameters, e.g., setting $\alpha = 1$, $w = 10$, and $l = 0$ generates a pattern more similar to Figure 4(b).

in particular, this paper formalizes early social psychology models to show that they predict the key anomalies from prospect theory.

Given the similar predictions of these models, distinguishing between the theories is not always necessary. If consumers and investors are afraid of looking foolish to friends and family, one modeling choice is to assume prospect theory or regret theory behavior. However, the predictions can vary depending on the information and incentive environments so distinguishing between the theories can sometimes be important. Loss aversion can be reversed when private information on skill allows for dare-taking behavior, and the pattern of probability weighting is reversed when observers are uninformed of the probabilities. These differences indicate when it is important to explicitly model the information flows, and also provide a basis for testing under what conditions the behavioral or informational approach to understanding risk is more predictive.

4 Appendix

Proof of Lemma 1: From the discussion in the text, if (5) is strictly positive at $\mu' = \Pr_F[s|win]$, or is weakly positive and approaching zero from above, it cannot be negative for any μ' since that would require crossing zero twice. Looking back to Figure 2(b), (5) evaluated at $\mu' = \Pr_F[s|win]$ equals the triangle under $\int_{\Pr_G[s|lose]}^{\Pr_G[s|win]} Q(\mu)d\mu$ plus the trapezoid below $\int_{\Pr_G[s|win]}^{\Pr_F[s|win]} Q(\mu)d\mu$ minus the triangle below $\int_{\Pr_F[s|lose]}^{\mu'} P(\mu)d\mu$, or

$$
\begin{aligned}
&\frac{1}{2}\int_0^{\Pr_G[s|win]} Q(\mu)d\mu \left(\Pr_G[s|win] - \Pr_G[s|lose]\right) \\
&+\frac{1}{2}\left(\int_0^{\Pr_G[s|win]} Q(\mu)d\mu + \int_0^{\Pr_F[s|win]} Q(\mu)d\mu\right)\left(\Pr_G[s|win] - \Pr_G[s|win]\right) \qquad (21)\\
&-\frac{1}{2}\int_0^{\Pr_F[s|win]} Q(\mu)d\mu \left(\Pr_F[s|win] - \Pr_F[s|lose]\right), \qquad (22)
\end{aligned}
$$

or

$$
\begin{aligned}
= \quad &\frac{1}{2}\Pr_G[lose]\left(\Pr_G[s|win] - \Pr_G[s|lose]\right)^2 \qquad (23)\\
&+\frac{1}{2}\left(\Pr_G[lose] + \left(\Pr_F[s|win] - \Pr_G[s|win]\right)\right)\left(\Pr_F[s|win] - \Pr_G[s|win]\right) \qquad (24)\\
&-\frac{1}{2}\Pr_F[win]\left(\Pr_F[s|win] - \Pr_F[s|lose]\right)^2, \qquad (25)
\end{aligned}
$$

which simplifies to

$$
\frac{\Delta^2}{2}\left(\Pr[s]\Pr[u]\right)^2 \frac{\left(\Pr_G[win] - \Pr_F[win]\right)\left(\Pr_F[win] + \Pr_G[win] - 1\right)}{\Pr_G[win]\Pr_F[win]\Pr_G[lose]\Pr_F[lose]}, \qquad (26)
$$

so $\Pr_F[win] + \Pr_G[win] > 1$, or $\Pr_G[win] > \Pr_F[lose]$, is sufficient for TOSD, i.e., the long-shot is less likely to lose than the sure thing is to win. If $\Pr_F[win] + \Pr_G[win] = 1$ so (5) equals zero at $\mu' = \Pr_F[s|win]$, then TOSD holds if the first derivative at $\mu' = \Pr_F[s|win]$ is zero and the second left derivative is positive, implying that (5) is positive in the left neighborhood of $\Pr_F[s|win]$. Checking, the first derivative is $\int_0^{\Pr_F[s|win]} (Q(\mu) - P(\mu))\, d\mu dt$ which is zero since $E_Q[\mu] = E_P[\mu] = \Pr[s]$ and there is no mass for either distribution after $\Pr_F[s|win]$ under the assumptions of $\Pr_F[win] < \Pr_G[win]$ and equal skill gaps. The second left derivative is $Q(\Pr_G[s|win]) - P(\Pr_G[s|win]) > 0$ so TOSD holds.

Proof of Proposition 2: (i) Generalizing the text's analysis of the separating equilibrium, let Δ_θ be type θ's mixed strategy probability of gambling. After observing that a gamble is refused, the observer believes it came from type g with probability ρ, and after observing that a gamble is accepted the observer believes it came from type g with probability γ. In the latter case this belief is then updated by the outcome x. Since our equilibrium concept is sequential equilibrium, beliefs on the equilibrium path are consistent with equilibrium strategies,

$$\rho = \Pr[g]\,(1-\Delta_g) \,/\, (\Pr[g]\,(1-\Delta_g) + \Pr[b]\,(1-\Delta_b))\,,$$

$$\gamma = \Pr[g]\Delta_g / \,(\Pr[g]\Delta_g + \Pr[b]\Delta_b)\,.$$

Given beliefs ρ and γ, the embarrassment premium from a gamble is

$$\pi_\theta \equiv v(\Pr_\rho[s]) - \Pr[win|\theta] v(\Pr_\gamma[s|win]) - \Pr[lose|\theta] v(\Pr_\gamma[s|lose]), \tag{27}$$

where $\Pr_\rho[s] = \rho \Pr[s|g] + (1-\rho)\Pr[s|b]$, and

$$\Pr_\gamma[s|x] = \Pr_\gamma[s] + \frac{\Pr[x|s] - \Pr[x|u]}{\Pr_\gamma[x]} \Pr_\gamma[s] \Pr_\gamma[u], \tag{28}$$

where $\Pr_\gamma[s] = \gamma \Pr[s|g] + (1-\gamma)\Pr[s|b]$ and $\Pr_\gamma[x] = \gamma \Pr[x|g] + (1-\gamma)\Pr[x|b]$.

A separating equilibrium exists if, for $\rho = 0$ and $\gamma = 1$, $\pi_g \le E[x|g] - z$ and $\pi_b \ge E[x|b] - z$. A partial pooling equilibrium where g always gambles and b mixes, $\Delta_g = 1$ and $\Delta_b \in (0,1)$, exists if for $\rho = 0$ and some $\gamma \in (\Pr[g], 1)$, $\pi_b = E[x|b] - z$ and $\pi_g \le E[x|g] - z$. Beliefs at the limit of $\Delta_b = \Delta_g = 1$ where both types gamble are off the equilibrium path, so for a sequential equilibrium these off path beliefs are the limiting beliefs of a mixed strategy equilibrium as the mixed strategies approach these pure strategies [Formalize more]. In the limit as $\Delta_b \to 1$, beliefs for who refuse hold constant at $\rho = 0$ as beliefs for who gambles converge to $\Pr[g]$. Hence a both gamble equilibrium exists if, for $\rho = 0$ and $\gamma = \Pr[g]$, $\pi_\theta \le E[x|\theta] - z$. A partial pooling equilibrium where b never gambles and g mixes, $\Delta_b = 0$ and $\Delta_g \in (0,1)$ exists if, for some $\rho \in (0, \Pr[g])$ and $\gamma = 1$, $\pi_b \ge E[x|b] - z$ and $\pi_g = E[x|g] - z$. In the limit as $\Delta_g \to 0$, beliefs for who refuses converge to $\rho = \Pr[g]$ and beliefs for who gambles stay constant at $\gamma = 1$, so a neither gamble equilibrium exists if, for $\rho = \Pr[g]$ and $\gamma = 1$, $\pi_\theta \ge E[x|\theta] - z$. Since $\pi_g > \pi_b$ for any candidate equilibrium beliefs, any of the above equilibria exist for some choice of z. For the same reason, other equilibria such as g refusing and b gambling cannot exist.

For the equilibrium embarrassment premia, we divide the above non-separating equilibria into two cases for proving (a),(b), and (c) of part (i) of the proposition: the b-mixes partial pooling equilibria and the both gamble equilibria where we focus on π_b since only b can be made indifferent in equilibrium, and the g-mixes partial pooling equilibria and the neither gamble equilibria where we focus on π_g since only g can be made indifferent in equilibrium.

In the former case $\rho = 0$ and $\gamma \in [\Pr[g], 1)$. (a) As $\Pr[s|g] - \Pr[s|b]$ goes to 0, $\Pr_\rho[s]$ goes to $\Pr[s]$, and $\Pr_\gamma[x]$ goes to $\Pr[s]$ so $\pi_b > 0$ just as in the no signal case of (**??**). (b) As $\Pr[s|g] - \Pr[s|b]$ goes to

1, $\Pr_\rho[s]$ goes to 0 since $\Pr[s|b]$ goes to 0, and $\Pr_\gamma[s|x]$ is strictly positive, so $\pi_b < 0$. (c) As the skill gap $\Pr[win|s] - \Pr[win|u]$ goes to 0, π_b goes to $v(\Pr[s|b]) - v(\Pr_\gamma[s]) < 0$.

In the latter case $\rho = [0, \Pr[g])$ and $\gamma = 1$. (a) As $\Pr[s|g] - \Pr[s|b]$ goes to 0, again $\Pr_\rho[s]$ goes to $\Pr[s]$, and $\Pr_\gamma[x]$ goes to $\Pr[s]$ so again $\pi_g > 0$ just as in the no signal case of (**??**). (b) As $\Pr[s|g] - \Pr[s|b]$ goes to 1, $\Pr_\rho[s]$ goes to ρ since $\Pr[s|b]$ goes to 0, and $\Pr_\gamma[s|x]$ goes to 1 since $\Pr[s|g]$ goes to 1, so π_g goes to $v(\rho) - v(1) < 0$. (c) As the skill gap $\Pr[win|s] - \Pr[win|u]$ goes to 0, π_g goes to $v(\Pr_\rho[s]) - v(\Pr[s]) < 0$.

(ii) Comparing a long-shot F with a complementary sure-thing G, the text establishes that $\pi_{G,g} > \pi_{F,g}$ in a separating equilibrium. To confirm that this implies $\pi_{G,b} > \pi_{F,b}$, note that for any gamble

$$\begin{aligned}
\pi_g - \pi_b &= (v(\Pr[s|b]) - E[v(\Pr[s|x,g])|g]) - (v(\Pr[s|b]) - E[v(\Pr[s|x,g])|b]) \\
&= (-\Pr[win|g]v(\Pr[s|win,g]) - \Pr[lose|g]\, v(\Pr[s|lose,g])) \\
&\quad - (-\Pr[win|b]v(\Pr[s|win,g]) - \Pr[lose|b]\, v(\Pr[s|lose,g])) \\
&= (-\Pr[win|g] + \Pr[win|b])\, v(\Pr[s|win,g]) \\
&\quad + (-\Pr[lose|g] + \Pr[lose|b])\, v(\Pr[s|lose,g]) \\
&= (-\Pr[win|g] + \Pr[win|b])\, v(\Pr[s|win,g]) \\
&\quad + (-1 + \Pr[win|g] + 1 - \Pr[win|b])\, v(\Pr[s|lose,g]) \\
&= (-\Pr[win|g] + \Pr[win|b])\, (v(\Pr[s|win,g]) - v(\Pr[s|lose,g])).
\end{aligned}$$

Therefore, since $\Pr_F[win|g] - \Pr_F[win|b] = \Pr_G[win|g] - \Pr_G[win|b]$,

$$\begin{aligned}
& (\pi_{G,g} - \pi_{F,g}) - (\pi_{G,b} - \pi_{F,b}) \\
= & -(\pi_{F,g} - \pi_{F,b}) + (\pi_{G,g} - \pi_{G,b}) \\
= & -\left(-\Pr_F[win|g] + \Pr_F[win|b]\right)\left(v(\Pr_F[s|win,g]) - v(\Pr_F[s|lose,g])\right) \\
& + \left(-\Pr_G[win|g] + \Pr_G[win|b]\right)\left(v(\Pr_G[s|win,g]) - v(\Pr_G[s|lose,g])\right) \\
\propto & \left(v(\Pr_F[s|win,g]) - v(\Pr_F[s|lose,g])\right) - \left(v(\Pr_G[s|win,g]) - v(\Pr_G[s|lose,g])\right) < 0
\end{aligned} \tag{29}$$

where the inequality follows from $v'' < 0$ since $\Pr[s|win,g] - \Pr[s|lose,g]$ is decreasing in $\Pr[win|g]$ for the same skill gap from (11).

Now consider again the two cases of g indifference and b indifference above. For the first case only g types gamble, so the embarrassment premium difference for g is

$$\pi_{G,g} - \pi_{F,g} = E[v(\Pr_G[s|x,g])|g] - E[v(\Pr_F[s|x,g])|g] > 0$$

where the inequality follows since the gap is same as for the separating equilibrium.

For the second case both types gamble. In order to apply Lemma 1then we can consider the embarrassment premium for a type that is a weighted average of g and b according to γ with corresponding $\Pr_\gamma[win] = \gamma \Pr[win|g] + (1-\gamma)\Pr[win|b]$. For such a weighted type, gambling over skill is a fair gamble that meets the conditions of Lemma 1 by the same argument in the text used for the separating equilibrium. Therefore, using the γ subscript to indicate this weighted type, $\pi_{G,\gamma} > \pi_{F,\gamma}$. To show that the embarrassment premium difference for b is worse than the average case of γ, note that for any gamble,

$$\pi_\gamma - \pi_b = \left(-\Pr_\gamma[win] + \Pr[win|b]\right)\left(v(\Pr_\gamma[s|win]) - v(\Pr_\gamma[s|lose])\right)$$

so the difference in embarrassment premia for the average type and the b type is $(\pi_{G,\gamma} - \pi_{F,\gamma}) - (\pi_{G,b} - \pi_{F,b})$ which equals

$$\begin{aligned}
& -(\pi_{F,\gamma} - \pi_{F,b}) + (\pi_{G,\gamma} - \pi_{G,b}) \\
= & -\left(-\Pr_{F,\gamma}[win] + \Pr_{F}[win|b]\right)\left(v(\Pr_{F,\gamma}[s|win]) - v(\Pr_{F,\gamma}[s|lose])\right) \\
& + \left(-\Pr_{G,\gamma}[win] + \Pr_{G}[win|b]\right)\left(v(\Pr_{G}[s|win,g]) - v(\Pr_{G}[s|lose,g])\right) \\
= & -\left(-\gamma \Pr_{F}[win|g] - (1-\gamma)\Pr_{F}[win|b] + \Pr_{F}[win|b]\right)\left(v(\Pr_{F,\gamma}[s|win]) - v(\Pr_{F,\gamma}[s|lose])\right) \\
& + \left(-\gamma \Pr_{G}[win|g] - (1-\gamma)\Pr_{G}[win|b] + \Pr_{G}[win|b]\right)\left(v(\Pr_{G}[s|win,g]) - v(\Pr_{G}[s|lose,g])\right) \\
= & \ \gamma\left(\Pr_{F}[win|g] - \Pr_{F}[win|g]\right) \times \\
& \left(\left(v(\Pr_{F}[s|win,g]) - v(\Pr_{F}[s|lose,g])\right) - \left(v(\Pr_{G}[s|win,g]) - v(\Pr_{G}[s|lose,g])\right)\right) < 0
\end{aligned}$$

where the inequality follows by the same argument as above. ■

Proof of Proposition 3: (i) Similar to the performance skill case, for given beliefs ρ and γ, the embarrassment premium from a gamble is

$$\pi_\theta \equiv v(\Pr[s]) - \Pr[win|\theta]v(\Pr_{\gamma}[s|win]) - \Pr[lose|\theta]v(\Pr_{\gamma}[s|lose]), \tag{30}$$

where updated skill is now

$$\Pr_{\gamma}[s|x] = \Pr[s] + \frac{\Pr_\gamma[x|s] - \Pr_\gamma[x|u]}{\Pr_\gamma[x]}\Pr[s]\Pr[u] \tag{31}$$

with $\Pr_\gamma[x|s] = \gamma\Pr[x|s,g] + (1-\gamma)\Pr[x|s,b]$ and $\Pr_\gamma[x|s] = \gamma\Pr[x|s,g] + (1-\gamma)\Pr[x|s,b]$. Different from performance skill, estimated evaluation skill from not gambling always equals the prior $\Pr[s]$ since θ alone provides no information. Also different is the skill gap's effect on updating disappears if gambling is believed to be by each type with equal probability, and is reversed if $\gamma < 1/2$. As noted in the text, we restrict attention to equilibria where the skill gap is non-negative, $\gamma \geq 1/2$. [If negative then positive risk premia possible for b, but still faces financial loss from losing. Note for $\Pr[g] < 1/2$ this precludes a full pooling equilibrium].

$$\begin{aligned}
\Pr_{\gamma}[win|s] &= \Pr[win|s,g]\Pr[g]\Delta_g + \Pr[win|s,b]\Pr[b]\Delta_b \\
\Pr_{\gamma}[win|u] &= \Pr[win|u,g]\Pr[g]\Delta_g + \Pr[win|u,b]\Pr[b]\Delta_b \\
\Pr_{\gamma}[win|s] - \Pr_{\gamma}[win|u] &= (\Pr[win|s,g] - \Pr[win|u,g])\Pr[g]\Delta_g + (\Pr[win|s,b] - \Pr[win|u,b])\Pr[b]\Delta_b \\
&= (\Pr[win|s,g] - \Pr[win|u,g])\Pr[g]\Delta_g + (\Pr[win|s,b] - \Pr[win|u,b])\Pr[b]\Delta_b \\
&> 0 \text{ if } \Pr[g]\Delta_g > \Pr[b]\Delta_b.
\end{aligned}$$

Restating (12) and (13), a separating equilibrium exists if, for $\gamma = 1$, $\pi_g \leq E[x|g] - z$ and $\pi_b \geq E[x|b] - z$. For g the expected skill estimate is the prior $\Pr[s]$ so the uncertainty implies $\pi_g > 0$, and the distribution is worse for b so $\pi_b > 0$. A g-mixes partial pooling equilibrium exists if, for $\gamma = 1$, $\pi_g = E[x|g] - z$ and $\pi_b \leq E[x|b] - z$ where again $\pi_g > 0$. In the limit as $\Delta_g \to 0$, beliefs for who refuses remain irrelevant and beliefs for who gambles stay constant at $\gamma = 1$, so a neither gamble equilibrium exists [formalize more] if, for $\gamma = 1$, $\pi_\theta \leq E[x|\theta] - z$ where again $\pi_g > 0$. A b-mixes partial pooling equilibrium exists if, for $\gamma \in (\Pr[g], 1)$, $\pi_g \geq E[x|g] - z$ and $\pi_b = E[x|b] - z$. The expected estimated skill for the belief-weighted mix of g and b is always the prior $\Pr[s]$, so estimated skill is worse for type b under the restriction that the skill gap remains positive, $\gamma \geq 1/2$, or $\Delta_b < \Pr[g]/\Pr[b]$. It is also more variable and also variable, so $\pi_b > 0$. A both-gamble equilibrium where $\gamma = \Pr[g]$ and the skill gap is positive if $\Pr[g] \geq 1/2$, for $\gamma = \Pr[g]$, $\pi_\theta \geq E[x|\theta] - z$. The expected estimated skill unconditional on θ is the prior $\Pr[s]$, so estimated skill is strictly worse for type b and more variable if $\gamma > 1/2$, while if $\gamma = 1/2$ there is no updating. So $\pi_b \geq 0$.

(ii) To extend the result in the text that $\pi_{G,g} > \pi_{F,g}$ to also show $\pi_{G,b} > \pi_{F,b}$, from (29) $(\pi_{G,g} - \pi_{F,g}) - (\pi_{G,b} - \pi_{F,b})$ has the same sign as

$$\Big(v(\Pr_F[s|win, g]) - v(\Pr_F[s|lose, g])\Big) - \Big(v(\Pr_G[s|win, g]) - v(\Pr_G[s|lose, g])\Big) > 0$$

where the inequality follows since $\Pr[s|win, g] - \Pr[s|lose, g]$ is decreasing in $\Pr[win|g]$ for the same skill gap from (12) and (13).■

References

[1] **Atkinson, John W.** 1957. "Motivational Determinants of Risk-Taking Behavior," *Psychological Review*, 64(6): 359–372.

[2] **Atkinson, John W., Jarvis R. Bastian, Robert W. Earl, and George H. Litwin.** 1960. "The Achievement Motive, Goal Setting, and Probability Preferences," *Journal of Abnormal and Social Psychology*, 60(1): 27–36.

[3] **Andreoni, James and William T. Harbaugh**. 2010. "Unexpected Utility: Experimental Tests of Five Key Questions about Preferences over Risk," working paper.

[4] **Bell, David E.** 1982. "Regret in Decision Making under Uncertainty," *Operations Research*, 30(5): 961–981.

[5] **Benabou, Roland and Jean Tirole.** 2002. "Self-Confidence and Personal Motivation," *Quarterly Journal of Economics*, 117(3): 871–915.

[6] **Burks, Stephen V., Jeffrey P. Carpenter, Lorenz Goette, and Aldo Rustichini**. 2013. "Overconfidence and Social Signalling," *Review of Economic Studies*, 80(3): 949–983.

[7] **Charness, Gary, Aldo Rustichini, and Jeroen Van de Ven**. 2018. "Self-confidence and Strategic Behavior," *Experimental Economics* 21(1): 72–98.

[8] **Chevalier, Judith and Glenn Ellison.** 1999. "Career Concerns of Mutual Fund Managers," *Quarterly Journal of Economics*, 114(2): 389–432.

[9] **Chen, Ying**. 2016. "Career Concerns and Excessive Risk Taking," J*ournal of Economics and Management Strategy* 24(1): 110–130.

[10] **Cowen, Tyler and Amihai Glazer.** 2007. "Esteem and Ignorance," *Journal of Economic Behavior and Organization*, 63(3): 373–383.

[11] **Chung, Kim Sau and Peter Eso**. 2013. "Persuasion and Learning by Countersignaling," *Economics Letters*, 121(3): 487–491.

[12] **Ebert, Sebastian and Daniel Wiesen**. 2011. "Testing for Prudence and Skewness Seeking," *Management Science* 57(7): 1334–1349.

[13] **Eckel, Catherine and Philip Grossman.** 2008. "Men, Women, and Risk Aversion: Experimental Evidence," 2008, *Handbook of Experimental Economic Results*, 1: 1061-1073. Elsevier.

[14] **Gneezy, Uri, John A. List, and George Wu**. 2006. "The Uncertainty Effect: When a Risky Prospect is Valued Less than its Worst Possible Outcome," *Quarterly Journal of Economics*, 121(4): 1283–1309.

[15] **Goffman, Erving.** 1959. *The Presentation of Self in Everyday Life*. Doubleday: New York, NY.

[16] **Gul, Faruk.** 1991. "A Theory of Disappointment Aversion," *Econometrica*, 59(3): 667–686.

[17] **Harbaugh, Richmond**. 1996. "Falling Behind the Joneses: Relative Consumption and the Growth-Savings Paradox," *Economics Letters*, 53(3): 297–304.

[18] **Harrison, Glenn W. and Don Ross.** 2017. "The Empirical Adequacy of Cumulative Prospect Theory and its Implications for Normative Assessment," *Journal of Economic Methodology*, 24(2): 150–165.

[19] **Holmstrom, Bengt**. 1982/1999. "Managerial Incentive Problems - A Dynamic Perspective," festschrift for Lars Wahlback (1982), republished in *Review of Economic Studies*, 66(1): 169–182.

[20] **Holmstrom, Bengt and Joan Ricart I Costa**. 1986. "Managerial Incentives and Capital Management," *Quarterly Journal of Economics*, 101(4): 835–860.

[21] **Holmstrom, Bengt**. 2016. "Pay For Performance and Beyond," Nobel Prize Lecture.

[22] **James, William**. 1890. *The Principles of Psychology*, Henry Holt and Co., New York.

[23] **Jones, Edward E. and Steven Berglas**. 1978. "Control of Attributions About the Self Through Self-Handicapping Strategies: The Appeal of Alcohol and the Role of Underachievement," *Personality and Social Psychology Bulletin*, 4(2): 200–206.

[24] **Kahneman, Daniel and Amos Tversky**. 1979. "Prospect Theory: An Analysis of Decision under Risk," *Econometrica,* 47(2): 263–292.

[25] **Kahneman, Daniel.** 2002. "Maps of Bounded Rationality: A Perspective on Intuitive Judgment and Choice," Nobel Prize Lecture.

[26] **Karlin, Samuel**. 1967. *Total Positivity*. Vol. 1. Stanford University Press, 1968.

[27] **Kimball, Miles S.** 1990. "Precautionary Saving in the Small and in the Large," *Econometrica*, 58(1): 53–73.

[28] **Kolditz, T.A. and R.M. Arkin**. 1982. "An Impression Management Interpretation of the Self-Handicapping Strategy," *Journal of Personality and Social Psychology*, 43(3): 492–502.

[29] **Lau, Morten, Hong Il Yoo, and Hongming Zhao**. 2019. "The Reflection Effect and Fourfold Pattern of Risk Attitudes: A Structural Econometric Analysis," working paper.

[30] **Laury, Susan K. and Charles A. Holt**. 2008. "Further Reflections on the Reflection Effect," *Research in Experimental Economics*, 12: 404–440.

[31] **Loomes, Graham and Robert Sugden**. 1982. "Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty," *Economic Journal*, 92: 805–824.

[32] **Menezes, C. F., Geiss, C., and Tressler, J**. 1980. "Increasing Downside Risk." *American Economic Review*, 70(5): 921–932.

[33] **Pratt, John W.** 1964. "Risk Aversion in the Small and in the Large," *Econometrica*, 32(1/2): 122–136.

[34] **Quiggin, John**. 1982. "A Theory of Anticipated Utility," *Journal of Economic Behavior and Organization*, 3(4): 323–343.

[35] **Rabin, Matthew.** 2000. "Risk Aversion and Expected Utility: A Calibration Theorem," *Econometrica*, 68(5): 1281–1292.

[36] **Ruggeri, Kai, Sonia Alí, Mari Louise Berge, Giulia Bertoldo, Ludvig D. Bjørndal, Anna Cortijos-Bernabeu, Clair Davison et al.** 2020. "Replicating Patterns of Prospect Theory for Decision Under Risk," *Nature Human Behaviour*, May 18, 1–12.

[37] **Schlaifer, Robert**. 1969. *Analysis of Decisions under Uncertainty*, McGraw-Hill.

[38] **Steiner, Jakub and Colin Stewart**. 2016. "Perceiving Prospects Properly," *American Economic Review*, 106(7): 1601–1631.

[39] **Tversky, Amos and Daniel Kahneman**. 1981. "The Framing of Decisions and the Psychology of Choice," *Science*, 211(4481): 453–458.

[40] **Tversky, Amos and Daniel Kahneman**. 1992. "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty*, 5(4): 297–323.

[41] **Weitzman, Martin L.** 1980. "The "Ratchet Principle" and Performance Incentives," *The Bell Journal of Economics*, 11(1): 302–330.

[42] **Whitmore, G. A.** 1970. "Third-Degree Stochastic Dominance," *American Economic Review*, 60(3): 457–459.